现代数学译丛 12

多尺度计算方法

——均匀化和平均化

〔英〕Grigorios A.Pavliotis Andrew M.Stuart 著

郑健龙 李友云 钱国平 译



现代数学译丛 12

多尺度计算方法

——均匀化和平均化

〔英〕Grigorios A. Pavliotis Andrew M. Stuart 著

郑健龙 李友云 钱国平 译

科学出版社

北京

版权登记号 01-2010-2337

内 容 简 介

本书针对各类具有多尺度特性的问题给出简化数学处理方法(平均化和均匀化), 该方法可用于求解偏微分方程、随机微分方程、常微分方程以及 Markov 链.

全书共分三部分, 第一部分为背景资料; 第二部分为扰动展开, 给出此类问题的共性; 第三部分阐述了一些证明扰动方法的理论. 每章结束部分的讨论和文献目录中均对本章的一些结论进行了推广和扩展, 并附上参考文献. 除第 1 章外, 所有章节均提供相应练习.

本书既可作为高等院校本科和研究生教材, 也可作为教师、工程技术人员和业余爱好者的自学用书.

图书在版编目(CIP)数据

多尺度计算方法: 均匀化和平均化/(英) Grigorios A. Pavliotis, Andrew M. Stuart 著; 郑健龙, 李友云, 钱国平译. —北京: 科学出版社, 2010

(现代数学译丛; 12)

ISBN 978-7-03-027512-7

I. 多… II. ①G… ②A… ③郑… ④李… ⑤钱… III. 微分方程解法 IV. O175.1

中国版本图书馆 CIP 数据核字(2010) 第 094921 号

责任编辑: 陈玉琢 / 责任校对: 陈玉凤
责任印制: 钱玉芬 / 封面设计: 王 浩

科 学 出 版 社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

骏 杰 印 刷 厂 印刷

科学出版社发行 各地新华书店经销

*

2010 年 6 月第 一 版 开本: B5(720×1000)
2010 年 6 月第一次印刷 印张: 18 1/2
印数: 1—3 000 字数: 358 000

定价: 56.00 元

(如有印装质量问题, 我社负责调换)

译者的话

多尺度方法是当代科学技术发展的前沿, 在数学科学、力学科学、生物科学、材料科学、化学科学等多种学科中运用越来越广泛. 而此类方法却由于受语言和专业术语的限制而不能够进行推广, 因此译者希望本书的翻译工作有利于多尺度方法在各个领域中的推广和发展.

本书介绍了各类具有多尺度特性的问题并给出了简化数学处理方法 (平均化和均匀化方法), 这些方法可用于求解偏微分方程、随机微分方程、常微分方程以及 Markov 链. 全书共分三部分, 第一部分为背景资料, 给出书中用到的分析方法、概率方法以及随机过程方面的理论知识, 并对常微分方程、Markov 链、随机微分方程以及偏微分方程进行简要介绍; 第二部分为扰动展开, 举例说明从平均化、均匀化方法到研究工程中的常微分方程、Markov 链、随机微分方程以及椭圆型、抛物线型、运输型偏微分方程所用到的思想, 并讨论了不变流型; 第三部分举例说明了用于证明扰动展开正确性的严格方法. 某些章节对一些特殊情形进行了讨论. 当然读者没有必要按照上面的顺序阅读本书, 建议先对本书第二部分的扰动展开问题进行了解, 然后根据需要参考第一部分的背景资料, 针对实际情况, 形成对所研究问题的近似处理, 有必要时再学习其他理论部分. 在每章结束部分的讨论和文献目录中均给出本章一些结论的推广和扩展, 并附上文献参考. 除第 1 章外, 所有章节均提供了相应的练习内容可供读者选择.

此书内容广泛、解释详尽, 书中列举的一些特殊素材对那些数学工作者、科技工作者以及经常使用该方法解决大量应用问题的工程师们具有特别的启发意义. 本书既可作为高等院校本科和研究生教材, 也可作为教师、工程技术人员和业余爱好者的自学用书.

本书的翻译工作及出版, 得到了国家自然科学基金项目 (50808026, 50778026) 及教育部归国人员启动基金, 湖南省教育厅科研项目 (06B008) 资助, 特此鸣谢. 限于水平, 不足之处在所难免, 恳切希望读者不吝惠于批评指正.

译　者

2010 年 5 月

目　录

第一部分　背　景

第二部分　扰动展开

第三部分　理　论

第1章 引　　言

1.1 概　　述

本章介绍了书中所要研究问题的类型以及要采用的研究方法. 1.2 节中包含的 4 个例子说明了在偏微分方程、确定性动力系统、随机动力系统中这一系列的相关问题; 1.3 节将讨论平均化和均匀化方法, 这些方法适用于解决奇异扰动线性微分方程, 本书第二部分将通过扰动展开式重点阐述其在材料中的应用; 1.4 节给出了本章引用到的参考文献和讨论.

2.2 节将对本章所用到的记号加以说明. 本章中将对 ODEs, PDEs 以及 SDEs 这些理论在概念上进行了介绍, 并在后面的章节中给出其更为详尽、全面的阐述, 而且也认为有必要在本章中对本书主题进行介绍, 因为这样可以减轻读者在建立数学模型时的负担.

1.2 启 发 例 子

本节叙述了 4 个例子, 这些例子阐述了在本书中需要研究的一系列问题. 第 1 个例子是在二阶线性一致椭圆型偏微分方程情况下的均匀化问题; 第 2 个例子阐述了在抛物型时变偏微分方程 (对流扩散方程) 情况下的相关思想; 通过双曲型偏微分方程和常微分方程之间的转化、特征线法、抛物型微分方程和随机微分方程之间的转换以及 Itô公式, 已经知道通过降低维数来解决线性偏微分方程的平均化和均匀化方法同样也适用于解决常微分方程和随机微分方程, 这方面的内容将在第 3 个例子中阐述; 关于动力学系统中的消元法将在第 4 个例子中阐述.

1.2.1 例 I：复合材料中的稳态热传导问题

为了引入均匀化的思想, 考虑复合材料的稳态热传导问题, 相对于材料的宏观尺度, 这种复合材料的特性变化得更快. 如果用 $\Omega \subset \mathbb{R}^d$ 表示材料的区域, 则区域尺寸定义为材料的宏观尺度 L(如: 可以通过定义 $\mathrm{vol}(\Omega) = L^d$ 定义 L). 另外, 定义非匀质性材料的特征长度为该类问题的微观尺度 ε, 假设当 $\varepsilon \ll 1$ 时, $L = O(1)$. 稳态导热现象可由下面的椭圆边值问题描述: 温度场设为 $u^\varepsilon(x)$,

$$-\nabla \cdot (A^\varepsilon \nabla u^\varepsilon) = f, \quad x \in \Omega, \tag{1.2.1a}$$

$$u^\varepsilon = 0, \quad x \in \partial\Omega. \tag{1.2.1b}$$

这里假定当 $\varepsilon > 0$ 时, A^ε 一致正定.

为了解决具体问题以及便于分析计算, 假设 $A^\varepsilon = A(x/\varepsilon)$ 和热导率张量矩阵 $A(y)$ 在所有 d 方向上都是 1 周期的并具有正定性. 均匀化理论的目的是为了研究 $u^\varepsilon(\varepsilon \to 0)$ 的极限. 特别地, 再找到 u^ε 的极限所满足的方程就更可取了. 从物理学角度来说, 极限 $\varepsilon \to 0$ 对应于不均匀性会逐渐消失. 这样目标就是用具有常系数 $\bar{A}$ 的有效均匀化材料来代替原始的、振荡系数为 $A(x/\varepsilon)$ 的高度非均质材料, 称这个过程为**均匀化**.

在第 12 章和第 19 章中, 对 $A(y), f(x)$ 和 Ω 作适当假设可以得到其均匀化方程为

$$-\nabla \cdot (\bar{A}\nabla u) = f, \quad x \in \Omega, \tag{1.2.2a}$$

$$u = 0, \quad x \in \partial\Omega. \tag{1.2.2b}$$

用如下公式表示常数**均匀化导率张量** $\bar{A}$:

$$\bar{A} = \int_{\mathbb{T}^d} A(y)\left(I + (\nabla\chi(y))^{\mathrm{T}}\right)\mathrm{d}y. \tag{1.2.3}$$

(一阶) 校正量 $\chi(y)$ 是一个向量场, 它是如下周期性**单胞问题**的解:

$$-\nabla_y \cdot \left(\nabla_y \chi A^{\mathrm{T}}\right) = \nabla_y \cdot A^{\mathrm{T}}, \quad y \in \mathbb{T}^d, \tag{1.2.4}$$

其中定义 $\mathbb{T}^d$ 为 d 维单胞, 并且偏微分方程 (1.2.4) 具有周期边界条件.

计算有效系数 $\bar{A}$ 包括在单胞体上求解偏微分方程以及计算 (1.2.3) 的积分. 因此, 找到一个均匀化解 u 需要解两个椭圆型偏微分方程: 周期性单胞问题 (1.2.4), 它允许通过 (1.2.3) 来构造 $\bar{A}$ 以及 Dirichlet 问题 (1.2.2). 重要的是这些椭圆方程并不依赖于小尺度 ε. 在这些例子中, 这两个椭圆型偏微分方程可以用显示格式进行求解, 甚至当不是这些例子时, 它们也符合严密的分析或者能获取直接的数值解; 因为其没有迅速变化的系数, 因此, 它的计算量远远少于直接求解 (1.2.1) 的计算量.

除了通过扰动展开推导均匀化方程, 还将证明在某种意义下, 当 $\varepsilon \to 0$ 时, (1.2.1) 的解 $u^\varepsilon(x)$ 收敛于均匀化方程 (1.2.2) 的解 $u(x)$. 在第 12 章中通过扰动展开将得到均匀化方程. 有关具有高速振荡系数的二次椭圆型偏微分方程的均匀化理论将在第 19 章中给出严格的证明.

1.2.2　例 II：对流扩散方程的均匀化

现在给出如何将均匀化的思想运用到进化偏微分方程这一类的抛物型方程中去. 考虑在一种不可压缩流中的化学浸泡, 如空气中的工业废物或水中的染料 ①.

① 事实上, 许多试验都是采用这种方法来使流体运动更加形象化.

假定工业废物或染料不影响流体的速率 $v(x,t)$, 随流体传输且服从分子碰撞, 其浓度场 $T(x,t)$ 服从下面的对流扩散方程:

$$\frac{\partial T}{\partial t}+v\cdot\nabla T=D\Delta T,\quad (x,t)\in\mathbb{R}^d\times(0,T), \tag{1.2.5a}$$

$$T=T_0,\quad (x,t)\in\mathbb{R}^d\times\{0\}, \tag{1.2.5b}$$

其中 $D>0$ 是分子扩散系数, 并且 Laplacian 项捕捉到在这个不移动扩散例子中工业废物或染料的传播. 第一导数项捕捉到了由流体运动引起的工业废物或染料的对流, 希望研究对流和扩散的相互作用.

假定流体速率是光滑、稳定的, 在各个方向是 1 周期的且流体不可压缩 (或者是无散), 则 $v(x,t)=-b(x)$, 其中 $\nabla\cdot b(x)=0$ ①.

注意到如果 T_0 是一个常数, 则 T 常数是方程的解. 设想 $T_0(x)=g(\varepsilon x)$ 时, 使得开始时浓度在空间中慢慢变化, 因此, 期望浓度只在大尺度和时间尺度上显著变化是有道理的. 如果 b 在单位立方体中均值为 0, 则均匀化技巧能找到改变尺度后的浓度场 $T(x/\varepsilon,t/\varepsilon^2)$, 使得浓度的一阶变量随 ε 趋于 0 而收敛于如下扩散方程的解 $\bar{T}$:

$$\frac{\partial \bar{T}}{\partial t}=\mathcal{K}:\nabla\nabla\bar{T},\quad (x,t)\in\mathbb{R}^d\times(0,T),$$

$$\bar{T}=g,\quad (x,t)\in\mathbb{R}^d\times\{0\}.$$

因此, 其有效行为是完全扩散的. 此处, $\mathcal{K}$ 为**有效扩散张量**,

$$\mathcal{K}=DI+\int_{\mathbb{T}^d}b(y)\otimes\chi(y)\mathrm{d}y.$$

向量校正场 $\chi(y)$ 是如下**单胞问题的解**:

$$-D\Delta_y\chi(y)-b(y)\cdot\nabla_y\chi(y)=b(y),\quad y\in\mathbb{T}^d,$$

并具有周期性.

例 I 中所作的论述同样适用于如下情形: 找到均质场需要解两个偏微分方程 (一个是具有周期边界条件的椭圆型偏微分方程, 另一个是抛物型偏微分方程). 由于该方程不受 ε 影响, 则易于分析和得到解析解或数值解. 此外, 在运用该均匀化思想解决类似问题时, 容易得到近似解并能发现误差. 在第 13 章中, 运用扰动展开可以得到均匀化方程. 对于具有快速振荡系数的抛物型偏微分方程的均匀化问题, 其严格的理论将在第 20 章中给出.

① 选择这种记号是为了与第 13 章和第 20 章中使用的记号一致.

1.2.3 例III：平均化、均匀化及动力系统

当速度场 $v(x,t)=-b(x)$ 是稳定的并具有周期性, 但在区域内其平均值不等于零时, 考虑方程 (1.2.5). 令 $x\to x/\varepsilon$ 和 $t\to t/\varepsilon^{\alpha}$, 方程 (1.2.5) 变为

$$\frac{1}{\varepsilon^{2-\alpha}}\frac{\partial T}{\partial t}-\frac{1}{\varepsilon}b^{\varepsilon}\cdot\nabla T=D\Delta T,\quad (x,t)\in\mathbb{R}^d\times(0,T), \tag{1.2.6a}$$

$$T=f,\quad (x,t)\in\mathbb{R}^d\times\{0\}, \tag{1.2.6b}$$

其中 $b^{\varepsilon}=b(x/\varepsilon)$. 当 $D=0$ 时, 选择 $\alpha=1$, 这时可用特征线法求解此方程; 特征线可通过求解如下常微分方程得到:

$$\frac{\mathrm{d}x^{\varepsilon}}{\mathrm{d}t}=b\left(\frac{x^{\varepsilon}}{\varepsilon}\right).$$

由于 $b(x)$ 具有周期性, x/ε 随周期尺度迅速变化, 为了消除迅速振荡因素, 通常会采用平均的思路, 这样可以看出在时变偏微分方程中快速消除尺度与常微分方程的平均化密切相关. 它们之间更深层次的联系将在第 14 章和第 21 章中给出. 显然, 当 b 的均值为零或不为零 (在某种意义上) 时, 要求在不同时间尺度 (当 a 的选取不同时) 下对 (1.2.6) 进行研究来观察有趣的动态行为, 这个问题在第 14 章中讨论.

当 $D>0$ 且向量场 b 的均值为零时, 通常选择 $a=2$. 方程 (1.2.6) 就是向后 Kolmogorov 方程, 这时需要求解随机微分方程①

$$\frac{\mathrm{d}x^{\varepsilon}}{\mathrm{d}t}=\frac{1}{\varepsilon}b\left(\frac{x^{\varepsilon}}{\varepsilon}\right)+\sqrt{2D}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{1.2.7}$$

其中 $W(t)$ 是 $\mathbb{R}^d$ 上的标准 Brownian 运动. 这意味着对于 $x^{\varepsilon}(0)=x$, (1.2.6) 中 $T^{\varepsilon}(x,t)$ 的解析式可记为

$$T(x,t)=\mathbb{E}\left(f\left(x^{\varepsilon}(t)\right)|x^{\varepsilon}(0)=x\right),$$

其中 $\mathbb{E}$ 为运用 Brownian 运动测度 (Wiener 测度) 求得的平均值. 对于上述问题, 同样能够消除快速变化量 x/ε. 如果 b 是周期的、无散的且平均等于 0, 则前面例子的结果表明随着 $\varepsilon\to 0$, $x^{\varepsilon}(t)$ 收敛于 $X(t)$, 其中 $X(t)$ 服从扩散系数为 $\sqrt{2\mathcal{K}}$ 的 Brownian 运动,

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\sqrt{2\mathcal{K}}\frac{\mathrm{d}W}{\mathrm{d}t}.$$

① 实际上, 此处是将平流扩散方程 (1.2.5) 看成一个平流和扩散量的密度场方程, 其从本质上讲就是向前 Kolmogorov(Fokker-Planck) 方程. 但是, 由于 $v(x,t)$ 是无散的, 所以只能根据平流项来区别 Kolmogorov 方程, 并且 (1.2.6a) 实际上是与随机微分方程 (1.2.7) 对应的向后 Kolmogorov 方程.

此外, 如果 $K \geqslant DI$(在矩阵意义下), 通过给出无散对流场, 这种扩散相对分子扩散得到了加强. 抛物型偏微分方程均匀化和随机微分方程均匀化之间的联系将在第 11 章和第 13 章进行讨论. 在第 18 章和第 20 章分别给出了严格的随机微分方程均匀化理论和抛物型偏微分方程均匀化理论.

1.2.4 例IV: 动力系统中降维

推导均匀化椭圆型偏微分方程和抛物型偏微分方程的方法也同样适用于推导常微分方程和随机微分方程系统中的平均化或均匀化. 这样做导致有效方程不包含小尺度参数 ε, 因此, 更加有利于获取数值解并进行分析. 如下形式的动力系统就是一个典型的例子:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\varepsilon} g(x,y).$$

在这种情况下, 当 $\varepsilon \ll 1$ 时, 就使得这里存在一个尺度分离, 也就可以消除 y 和找到控制进化量 x 的近似方程

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X).$$

通过一个**不变流形**来消除 y 是最简单的情况, 情况更加复杂时是通过**平均**来消除 y 的.

在一些 $F \equiv 0$ 的情况下, 有必要延长方程的时间尺度: $t \to t/\varepsilon$, 以便于找到其中的非零影响, 则开始点的方程为

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{\varepsilon} f(x,y),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\varepsilon^2} g(x,y).$$

在这种情况下, 当消除 y 并且拉伸 x 的尺度来得到满足 x 的方程, 从而找到 f 的起伏量时, f 的平均必须等于 0. 如果 y 的行为比较强, 则近似方程可采用如下形式:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) + A(X)\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中 W 是标准单元内的 Brownian 运动. 这种对方程的推导过程称为**均匀化**.

在第 8, 10, 11 章中详述了基于这种思想的扰动展开. 至于线性偏微分方程, 运用分析技巧证明了原始方程与简化方程之间的误差估计, 将在第 15, 17, 18 章中给出严格的结论. 类似常微分方程和随机微分方程的 Markov 链的平均化结论将在第 9 章和第 16 章中给出.

1.3　平均化对均匀化

本书中, 为了获取线性算子方程的扰动展开, 根据统一性原理采用如下形式:

$$\mathcal{L}^{\varepsilon}u^{\varepsilon}=f \tag{1.3.1}$$

或

$$\frac{\partial u^{\varepsilon}}{\partial t}=\mathcal{L}^{\varepsilon}u^{\varepsilon}. \tag{1.3.2}$$

特别地, 将对如下形式的算子 $\mathcal{L}^{\varepsilon}$ 更感兴趣:

$$\mathcal{L}^{\varepsilon}=\frac{1}{\varepsilon}\mathcal{L}_0+\mathcal{L}_1 \tag{1.3.3}$$

或

$$\mathcal{L}^{\varepsilon}=\frac{1}{\varepsilon^2}\mathcal{L}_0+\frac{1}{\varepsilon}\mathcal{L}_1+\mathcal{L}_2. \tag{1.3.4}$$

在上述两个例子中, 假定 $\mathcal{L}_0$ 是非零空间, 并且在它的子空间中能捕捉到解的形态. 例如, (1.3.3), 这个例子就是**平均化**或是**一次扰动理论**, 它给出的是大数定理的一种形式. 第二个例子 (1.3.4) 就是**均匀化**或是**二次扰动理论**, 它给出的是中心极限定理的一种形式.

下面来看进化方程 (1.3.2). 这个例子将有助于研究抛物型偏微分方程和双曲型偏微分方程的平均化和均匀化, 并且借助于随机微分方程 Kolmogorov 方程和变异型也有利于研究常微分方程、Markov 链以及随机微分方程的平均化和均匀化 (Markov 链的向前方程和常微分方程的特征线方法). 同时也会研究在椭圆型偏微分方程均匀化问题中出现的方程 (1.3.1). 在这个例子中, 将会出现二次时间导数 (波动方程), 在这里叙述不是很详细, 但是发展的技巧的确可以运用, 并会给出这一类问题的参考文献. 当实施这一类展开式时, 重点阐述平均化和均匀化之间的区别以及证明这个 Fredholm 性质的重要性. 本书中也把这类方法运用到线性常微分方程的奇异扰动系统中, 不过只是简单地介绍了一些关键思想. 接下来的两小节说明了在第二部分中还要重复阐述的扰动展开.

1.3.1　线性系统的平均化

考虑如下线性常微分系统:

$$\frac{\mathrm{d}u^{\varepsilon}}{\mathrm{d}t}=\mathcal{L}^{\varepsilon}u^{\varepsilon}, \tag{1.3.5a}$$

$$\mathcal{L}^{\varepsilon}=\frac{1}{\varepsilon}\mathcal{L}_0+\mathcal{L}_1. \tag{1.3.5b}$$

在这里, 为了简单考虑有限维的例子, 此处 $\mathcal{L}_i \in \mathbb{R}^{d\times d}\,(i=0,1)$. 假定 $\mathcal{L}_0$ 的零空间 $\mathcal{N}(\mathcal{L}_0)$ 是由 ϕ 张成的一维空间, $\mathcal{N}(\mathcal{L}_0^{\mathrm{T}})$ 是由 ψ 张成的, 并且 $\langle\phi,\psi\rangle \neq 0$.

如果找到如下形式的解:

$$u^\varepsilon = u_0 + \varepsilon u_1 + O\left(\varepsilon^2\right),$$

则

$$\frac{1}{\varepsilon}\mathcal{L}_0 u_0 + \left(\mathcal{L}_0 u_1 + \mathcal{L}_1 u_0 - \frac{\mathrm{d}u_0}{\mathrm{d}t}\right) + O(\varepsilon) = 0.$$

考虑方程两边 ε 各次幂的系数可得

$$O(1/\varepsilon): \mathcal{L}_0 u_0 = 0;$$

$$O(1): \mathcal{L}_0 u_1 = \frac{\mathrm{d}u_0}{\mathrm{d}t} - \mathcal{L}_1 u_0.$$

第一个方程暗示了对于某一个 $\alpha = \alpha(t) \in \mathbb{R}$ 有 $u_0 = \alpha\phi$. 对第二个方程, 运用 Fredholm 性质可得 u_1 存在的充要条件为

$$\left\langle \psi, \frac{\mathrm{d}u_0}{\mathrm{d}t} - \mathcal{L}_1 u_0 \right\rangle = 0,$$

其中用 $\langle\cdot,\cdot\rangle$ 表示 $\mathbb{R}^d$ 上的内积. 代入 $u_0 = \alpha\phi$ 可推出幅值 α 满足如下方程:

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{\langle\psi, \mathcal{L}_1\phi\rangle}{\langle\psi,\phi\rangle}\alpha. \tag{1.3.6}$$

如果

$$\langle\psi, \mathcal{L}_1\phi\rangle \neq 0, \tag{1.3.7}$$

则方程 (1.3.6) 有非零解. 这样就找到了如下近似解:

$$u^\varepsilon(t) \approx u_0(t) = \alpha(t)\phi,$$

其中 $\alpha(t)$ 是方程 (1.3.6) 的解. 这个推导举例说明了将在第 9, 10, 14 章中用到的平均化方法. 注意到幅值 $\alpha(t)$ 既可按指数增长也可按指数衰减. 方程 (1.3.5a) 的解 $u^\varepsilon(t)\,(\varepsilon \ll 1)$ 的行为是由 $\alpha(t)$ 的行为决定的. 这就是本书中给出的此类方法的第一个例子, 可以通过研究更加简单的方程 (1.3.6) 来决定原始复杂问题 (1.3.5a) 的解的形态.

1.3.2 线性系统的均匀化

如果不满足 (1.3.7), 则均匀化方程变为

$$u_0 = \text{const}.$$

为了观察到有趣的、非平凡的动力系统, 有必要拉伸尺度, 即 $t \to t/\varepsilon$, 并寻找 (1.3.5a) 的解, 这就得到如下线性方程系统:

$$\frac{\mathrm{d}u^\varepsilon}{\mathrm{d}t} = \mathcal{L}^\varepsilon u^\varepsilon,$$

其中

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon^2}\mathcal{L}_0 + \frac{1}{\varepsilon}\mathcal{L}_1.$$

此处 $\mathcal{L}_i = \mathbb{R}^{d\times d}\,(i=0,1)$. 这样做不会失去广义性, 研究如下系统:

$$\frac{\mathrm{d}u^\varepsilon}{\mathrm{d}t} = \mathcal{L}^\varepsilon u^\varepsilon, \tag{1.3.8a}$$

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon^2}\mathcal{L}_0 + \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_2. \tag{1.3.8b}$$

其中 $\mathcal{L}_i = \mathbb{R}^{d\times d}\,(i=0,1,2)$. 像以前一样, 对 $\mathcal{L}_0$ 和 $\mathcal{L}_0^{\mathrm{T}}$ 的零空间作同样的假设, 并且还假设有如下**中心条件**:

$$\langle \psi, \mathcal{L}_1\phi\rangle = 0 \tag{1.3.9}$$

如果 (1.3.8a) 的解有如下形式:

$$u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + O\left(\varepsilon^3\right),$$

则

$$\frac{1}{\varepsilon^2}\mathcal{L}_0 u_0 + \frac{1}{\varepsilon}\left(\mathcal{L}_0 u_1 + \mathcal{L}_1 u_0\right) + \left(\mathcal{L}_0 u_2 + \mathcal{L}_1 u_1 + \mathcal{L}_2 u_0 - \frac{\mathrm{d}u_0}{\mathrm{d}t}\right) + O(\varepsilon) = 0.$$

令上述方程中 ε 的各次幂的系数为 0, 则可得如下方程:

$$O(1/\varepsilon^2) : \mathcal{L}_0 u_0 = 0;$$

$$O(1/\varepsilon) : \mathcal{L}_0 u_1 = -\mathcal{L}_1 u_0;$$

$$O(1) : \mathcal{L}_0 u_2 = \frac{\mathrm{d}u_0}{\mathrm{d}t} - \mathcal{L}_1 u_1 - \mathcal{L}_2 u_0.$$

第一个方程暗示了 $u_0 = \alpha\phi\,(\alpha = \alpha(t) \in \mathbb{R})$, 通过 (1.3.9) 和 Fredholm 变换可以推出第二个方程是可解的, 所以记

$$u_1 = -\alpha\eta, \quad \mathcal{L}_0\eta = \mathcal{L}_1\phi.$$

关于 η 的方程称为**单胞问题**, 通过使 η 与 ϕ 正交可以有选择地获得解的唯一性, 但是另外一种标准化过程将影响 u_0 近似解的格式.

至于第三个方程, 同样运用 Fredholm 变换可以得到

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{\langle \psi, \mathcal{L}_2\phi - \mathcal{L}_1\eta\rangle}{\langle \psi, \phi\rangle}\alpha. \tag{1.3.10}$$

这样已经找到了近似解

$$u^{\varepsilon}(t) \approx u_0(t) = \alpha(t)\phi.$$

注意到因为 (1.3.9), 则对 η 标准化过程的选择并不影响 u_0 的最终形式, 即通过增加 ϕ 和 η 的乘积不能影响 u_0 的格式. 这一种推导举例说明了将在第 11~13 章中用到的均匀化方法.

1.4 讨论和参考

已经对均匀化方法中的多尺度展开进行了系统的研究, 并且用于解决了偏微分方程理论中不同的问题[33], 并在文献 [241, 240, 246] 中可以找到这些思想在解决随机微分方程方面时更为深入的研究; 有关平均化和均匀化的典范性著作见文献 [27, 66, 153, 279]; 在随机过程, 特别是随机微分方程中, 平均化和均匀化方法的发展见文献 [111]; 随机微分方程的均匀化解法可见文献 [186, 238].

通过 Kolmogorov 以及 Fokker-Planck 公式来建立抛物型和椭圆型偏微分方程与随机微分方程之间的关系是贯穿本书的普遍思想, 见文献 [117, 236]. 自从 19 世纪 60 年代初以来, 又再次建立了偏微分方程中均匀化、平均化的思想与 Markov 随机过程极限理论之间的联系并进行了广泛的研究[110,165]; 随机过程中关于均匀化、平均化的更一般的理论在 19 世纪 70 年代得到了发展[181,240,241,246], 相关的数学理论专著见文献 [94, 168, 261].

在应用数学中, 无论是随机系统还是确定性系统, 消元法的应用已经成为非常重要的研究领域, 应用范围包括大气海洋科学[202,205], 分子动力学[284,285] 以及金融数学[106]; 随机以及确定系统消减模式各方面的问题, 包括其中的数值计算见文献 [125, 148, 150]; 更早的关于消减模式方面的研究见文献 [321].

多尺度方法可以对不同领域进行广泛的研究, 而绝不仅仅包含本书中所介绍的平均化和均匀化思想; 关于平均化、均匀化在常微分方程、随机微分方程以及偏微分方程中应用的一些重要主题在本书中不予讨论, 在这里, 仅是罗列了这些主题并给出相应的参考文献.

开发高效的数值解法来解决具有特征长度和时间尺度的问题是多尺度方法中一个非常重要的方面, 在文献 [83] 或者 [82, 119] 中可以找到有关多尺度数值解法的全面综述.

本书不包括确定的、无周期的均匀化方法, 有兴趣的读者可以参见文献 [66, ch. 13; 75; 311] 及其参考文献. 同样, 在这里不会讨论随机偏微分方程 (如边界条件是迅速振荡随机域的偏微分方程), 见文献 [177, 247]; 用顺向方法求解多尺度随机系统也可见文献 [34].

均匀化理论的一个主要应用方向是材料科学, 关于均匀化在弹力理论以及材料科学中的应用见文献 [7, 27, 67, 237]; 运用各种多尺度及均匀化方法对复合材料进行研究, 并对复合材料中有关均匀化系数边界问题以及相关模型的详细讨论见文献 [229]; 应用本书介绍的方法来解决化学工程中的问题也可见文献 [265, 325].

本书中仅仅是简略地提到了非线性偏微分方程中均匀化方法的使用, 有兴趣的读者可以参见论著 [95] 及其相关文献; 应用变分法非常适合解决非线性椭圆型偏微分方程, 此类相关问题及 Γ 收敛理论见文献 [48, 49, 72, 214]; 诸如 H 收敛及 G 收敛等相关概念在这里就不再展开论述了, 这方面的文献见文献 [296, 297, 310, 311].

在每章结尾处的讨论与参考中将会给出更为详细的参考文献.

第一部分　背　　景

第2章　分　　析

2.1　结　　构

为了后面分析的需要, 本章给出了一些需要用到的定义和定理 (大部分没有证明), 至于证明可以在 2.7 节中找到. 本节的介绍简洁明了, 因此, 不需要完全弄清楚这些证明, 而只是根据后面章节的需要进行参考. 主要是在第 15~21 章 (第三部分), 要求有严格的发展背景, 证明了一些关于动态问题维数递减的结论, 也有对偏微分方程平均化和均匀化结论的分析. 在第 8~14 章 (第二部分) 中, 通过微扰扩展, 推导出平均化结论和均匀化结论. 在本部分中, 许多材料是不需要的, 然而, 记号介绍和 Fredholm 变换将会贯穿整本书, 尤其是在第二部分.

微分方程渐近问题严格研究的自然背景是 Banach 空间和 Hilbert 空间中的弱收敛问题; 在本章中也将发展一种适当的弱收敛, 也就是双尺度收敛, 这对涉及周期均匀化的问题非常有用; 一个关于概率的补充章节 (即第 3 章) 对随机微分方程和 Markov 链的研究提供了一个适当的分析工具, 尤其对概率测度的弱收敛发展提供了一种背景材料.

2.2 节描述了贯穿本书大部分内容的符号, 尤其是微分算子的无索引符号, 如散度、梯度和拉普拉斯算子; Banach 和 Hilbert 空间理论的综述见 2.3 节; 2.4 节描述了函数空间的具体实例, 都是非常有用的; 2.5 节介绍了双尺度收敛的概念, 并强调了它对高速震荡周期问题研究的重要性. 本章以对 Hilbert 空间上线性方程的两个基本研究方法: Lax-Milgram 定理和 Fredholm 变换的讨论结束.

2.2　记　　号

在本书中将遇到标量、向量和矩阵场. 接下来的求和是针对索引集 $\{1,2,\cdots,d\}$, d 是空间维数. 用 $\mathbb{R}^d$ 表示 d 维欧氏空间, 用 $\mathbb{T}^d$ 表示 d 维单位, 它是定义在 $\mathbb{R}^d$ 中的单位立方体.

用 $\{e_i\}_{i=1}^d$ 表示 $\mathbb{R}^d$ 空间的标准基, 这样对每一个 $\xi \in \mathbb{R}^d$ 都可以表示为 $\xi = \sum_i \xi_i e_i$, 其中 $\xi_i = \langle \xi, e_i \rangle$, $\langle \cdot, \cdot \rangle$ 表示 $\mathbb{R}^d$ 空间中的标准内积. 用 $\cdot$ 表示两个向量的内积, 即

$$a \cdot b = \sum_i a_i b_i.$$

由此内积所导出来的欧氏范数, 可表示为

$$|a| = \sqrt{a \cdot a},$$

进一步可得

$$|a|^2 = \sum_i a_i^2, \quad a \in \mathbb{R}^d.$$

用 $|\cdot|_p$ 表示 $\mathbb{R}^d$ 空间上的 $\ell^p (1 \leqslant p \leqslant \infty)$ 范数, 记作 $|\cdot| = |\cdot|_2$. 另外, 对于算子 $A : \mathbb{R}^d \to \mathbb{R}^m$, 使用 $|\cdot|_p$ 表示算子的相容性范数, 它可定义为

$$|A|_p = \sup_{x \neq 0} \frac{|Ax|_p}{|x|_p}.$$

特别地, 使用记号 $|A| = |A|_2$ 表示由欧氏向量范数导出的算子范数.

矩阵之间的内积可以表示为

$$A : B = \mathrm{tr}(A^{\mathrm{T}} B) = \sum_{ij} a_{ij} b_{ij}.$$

由此内积所导出的范数称为 F 范数,

$$|A|_F = \sqrt{\mathrm{tr}\,(A^{\mathrm{T}} A)}. \tag{2.2.1}$$

需要指出的是, 任一对称矩阵 T 和任一矩阵 S 可满足

$$S : T = S^{\mathrm{T}} : T = \frac{1}{2} \left(S + S^{\mathrm{T}} \right) : T. \tag{2.2.2}$$

两个向量 a, b 之间的向量积可以用矩阵 $a \otimes b$ 表示, 并且由下式定义:

$$(a \otimes b) c = (b \cdot c) a, \quad \forall c \in \mathbb{R}^d.$$

更一般地, 对 $A, B \in \mathbb{R}^{m \times d}$, $A \otimes B \in \mathbb{R}^{m \times m}$ 满足

$$(A \otimes B) c = A B^{\mathrm{T}} c, \quad \forall c \in \mathbb{R}^m.$$

用 ∇ 和 $\nabla \cdot$ 表示 $\mathbb{R}^d$ 中的梯度和散度. 梯度把标量 (向量) 提升为向量 (矩阵), 而散度把向量 (矩阵) 压缩为标量 (向量). 梯度针对标量值函数 $\phi(z)$ 或向量值函数 $v(z)$, 即

$$(\nabla \phi)_i = \frac{\partial \phi}{\partial z_i}, \quad (\nabla v)_{ij} = \frac{\partial v_i}{\partial z_j}.$$

散度针对向量值函数和矩阵值函数, 即

$$\nabla \cdot v = \sum_i \frac{\partial v_i}{\partial z_i}, \quad (\nabla \cdot A)_i = \sum_j \frac{\partial A_{ij}}{\partial z_j}.$$

给定向量场 a,v, 使用记号

$$a \cdot \nabla v := (\nabla v)\, a.$$

这样通过计算向量 v 中的每一个分量 $a \cdot \nabla v_k$ 定义了总量. 同样, 也可以扩展到如下记号:

$$a \cdot \nabla \Theta,$$

其中 Θ 是张量①场, 通过定义分量形式.

由于梯度是针对标量和向量来定义的, 对于任意的标量 ϕ, 很容易理解如下表示:

$$\nabla\nabla\phi,$$

它是 ϕ 的 Hessian 矩阵, 其分量可表示为 $\partial^2\phi/\partial x_i \partial x_j$. 同样, 通过对向量 v 中的每一个标量分量运用 $\nabla\nabla$, 如下表示有意义:

$$\nabla\nabla v,$$

或者

$$\nabla\nabla\Theta$$

的确也是分量形式. 定义一个标量或者一个向量场的 Laplacian 算子如下:

$$\Delta\phi = \nabla \cdot \nabla\phi, \quad \Delta v = \nabla \cdot \nabla v.$$

由此可推出 $\Delta\phi = I : \nabla\nabla\phi$(见练习第 1 题). 通过运用分量的光滑性可得到 $\Delta\Theta$ 的定义.

在许多例子中, 有必要区分不同独立变量的梯度和散度, 使用 ∇_z 表示关于 z 坐标的梯度和散度. 同样, 也可以针对其他的独立变量. 当独立变量的选择很清晰时, 保留记号 ∇, 不用下标. 相似的记号可以使用于其他的算子, 如 Laplacian 算子.

2.3 Banach 空间和 Hilbert 空间

假设读者已经熟悉范数的定义、内积、向量、矩阵和内积空间. 接下来, 用 X 表示赋范向量空间, 并且用 $\|\cdot\|$ 表示其范数. 一个序列 $\{x_j\}_{j=1}^{\infty} \subset X$ 强收敛于 $x \in X$, 即如果

$$\lim_{j\to\infty} \|x_j - x\| = 0,$$

可写为

$$x_j \to x,$$

① 在正文中用张量表示二阶张量, 或平方矩阵.

而且如果 $\forall\varepsilon > 0$, 存在一个 $N \in \mathbb{N}$, 满足对所有的 $j, k \geqslant N$ 都有

$$\|x_j - x_k\| < \varepsilon,$$

就称 $\{x_j\}_{j=1}^{\infty} \subset X$ 是一个 Cauchy 序列.

每一个赋范空间 X 中的收敛序列都是 Cauchy 序列; 反之, 不一定成立. 然而, 如果 X 中的每一个 Cauchy 序列都是收敛的, 则称空间 X 是完备的.

2.3.1 Banach 空间

定义 2.1 Banach 空间 X 是完备的赋范向量空间.

定义 2.2 如果 X 是一个具有范数 $\|\cdot\|$ 的 Banach 空间, 并且映射 $\ell : X \to \mathbb{R}$ 满足

(i) $\ell(\alpha x + \beta y) = \alpha\ell(x) + \beta\ell(y), \forall x, y \in X, \alpha, \beta \in \mathbb{R}$;

(ii) $\exists C > 0, |\ell(x)| \leqslant C\|x\|, \forall x \in X$,

则称映射 ℓ 是 X 中的有界线性泛函.

定义 2.3 所有作用于 Banach 空间 X 中的有界线性泛函组成的集合称为 X 的对偶空间, 用 X^* 表示.

定理 2.4 如果 Banach 空间 X 的对偶空间为 X^*, 定义范数

$$\|\ell\| = \sup_{x \neq 0} \frac{|\ell(x)|}{\|x\|},$$

则对偶空间 X^* 是一个 Banach 空间.

定义 2.5 设 X 是一个 Banach 空间, 那么它的对偶空间的对偶空间与 X 等价, 即

$$(X^*)^* = X.$$

对偶空间的概念能使得介绍 X 空间中的另外一种拓扑, 即所谓的弱拓扑.

定义 2.6 序列 $\{x_n\}_{n=1}^{\infty}$ 弱收敛于 $x \in X$, 即如果

$$\ell(x_n) \to \ell(x), \quad \forall \ell \in X^*,$$

可写成如下形式:

$$x_n \rightharpoonup x.$$

每一个强收敛序列也是弱收敛, 然而逆命题不成立. 弱收敛的重要性源于如下定理, 特别是 (ii):

定理 2.7 设 X 是 Banach 空间,

(i) X 中每一个弱收敛序列都是有界的;

(ii) **(Eberlein-Smuljan)** 假定 X 是自反的, 则可以从 X 中的每一个有界序列中抽出一个弱收敛的子序列.

除了 X 中的弱拓扑, X 和 X^* 的对偶性能够使得定义 X^* 中的一个拓扑.

定义 2.8 设 X 是 Banach 空间, 如果一个序列 $\{\ell_n\}_{n=1}^{\infty} \subset X^*$ 满足

$$\lim_{n\to\infty} \ell_n(x) = \ell(x), \quad \forall x \in X,$$

则称序列 $\{\ell_n\}_{n=1}^{\infty}$ 弱 $*$ 收敛于 $\ell \in X^*$, 可写为

$$\ell_n \overset{*}{\rightharpoonup} \ell.$$

评论: 如果 X 是自反的, 则弱 $*$ 收敛等同于弱收敛.

如果 X 是可分的, 类似于定理 2.7(ii) 的紧性结果将在 X^* 中的有界序列中体现, 但是没有自反性. 为了定义这一个概念, 回顾 X 的子集 X_0 是稠密的, 即对每一个 $x \in X$, 存在一个序列 $\{x_j\}_{j=1}^{\infty} \subset X_0$ 收敛于 x, 也就是说, 如果 X_0 的闭包是 X, 即 $\bar{X}_0 = X$, 则称 X_0 在 X 中稠密.

定义 2.9 如果 Banach 空间 X 包含一个可数的稠密子集, 则 X 是可分的.

X^* 中的序列具有紧性定理, 其说明如下:

定理 2.10 设 X 是可分的 Banach 空间, 则 X^* 中的任一有界序列可抽出一个弱 $*$ 收敛子列.

X^* 中的每一个弱收敛序列是弱 $*$ 收敛, 但是逆命题不真, 除非 X 是自反的.

2.3.2 Hilbert 空间

定义 2.11 完备的内积空间是 Hilbert 空间.

用 $(\cdot,\cdot)$ 表示内积, 并且显然每一个 Hilbert 空间是一个在 H 上具有如下范数的 Banach 空间:

$$\|x\| = (x,x)^{1/2},$$

而且 H 中的所有元素都满足**Cauchy-Schwarz 不等式**

$$|(u,v)| \leqslant \|u\|\,\|v\|.$$

Hilbert 空间 H 的一个非常重要的性质就是能够通过 Riesz 表示定理, 找到 H 的对偶空间是它本身.

定理 2.12 (Riesz 表示) 对于每一个 $\ell \in H^*$, 存在一个唯一的 $y \in H$ 满足

$$\ell(x) = (x,y), \quad \forall x \in H.$$

经常用[1]$\langle\cdot,\cdot\rangle_{H^*,H}$ 表示 ℓ 与 $x\in H$ 之间的作用, 称为一个对偶对,

$$\langle \ell, x\rangle_{H^*,H} := \ell(x).$$

Riesz 表示定理暗含了每一个 Hilbert 空间都是自反的, 也暗示了定理 2.7, 而且弱收敛的定义可简化为

$$x_n \rightharpoonup x \Leftrightarrow (x_n - x, y) \to 0, \quad \forall y \in H.$$

2.4 函数空间

2.4.1 连续函数空间

设 Ω 是 $\mathbb{R}^d$ 中的一个开集. 用 $C(\bar\Omega)$ 表示连续函数空间, 定义 $f:\bar\Omega\to\mathbb{R}$ 是 $C(\bar\Omega)$ 中的连续函数, 在 $C(\bar\Omega)$ 中定义如下最大模范数:

$$\|f\|_{C(\bar\Omega)} = \sup_{x\in\bar\Omega}|f(x)|,$$

则 $C(\bar\Omega)$ 是一个 Banach 空间. 类似地, 可以定义 k 阶连续导数空间 $C^k(\bar\Omega)$ 和无穷光滑函数空间 $C^\infty(\bar\Omega)$.

$C_0^k(\bar\Omega)$ 表示 $\bar\Omega$ 中具有紧支集 k 阶连续微分函数的集合所构成的空间. $C_0^\infty(\bar\Omega)$ 表示 $\bar\Omega$ 中具有紧支集无穷光滑函数的空间. $C_b^k(\mathbb{R}^d)$ 表示定义在 $\mathbb{R}^d$ 到 $\mathbb{R}$ 的 k 阶有界连续微分函数的集合所构成的空间. $C_b^\infty(\mathbb{R}^d)$ 表示定义在 $\mathbb{R}^d$ 到 $\mathbb{R}$ 的有界无穷光滑函数的集合所构成的空间. 更一般地, 设 (E,ρ) 表示度量空间, 则 $C_b(E)$ 表示 (E,ρ) 上的有界连续函数空间. 也可以把先前所有函数 $f:\Omega\to X(X=\mathbb{R}^d$ 或 $\mathbb{R}^{d\times d})$ 扩展到上述度量空间. 在所有这些例子中, 当意思很明了时, 通常简写为 $C(\bar\Omega)$ 或 $C(\bar\Omega;\mathrm{X})$. 至于更广义的 Banach 空间值函数 f, 见定义 2.22.

2.4.2 L^p 空间

设 $1\leqslant p\leqslant\infty$, $f:\Omega\to\mathbb{R}$ 表示 Lebesgue 测度函数 (见第 3 章的测度定义), 并且, 定义 L^p 范数如下:

$$\|f\|_{L^p(\Omega)} := \begin{cases} \left(\int_\Omega |f|^p\mathrm{d}x\right)^{1/p}, & 1\leqslant p<\infty, \\ \operatorname*{ess\,sup}_{\Omega}|f|, & p=\infty. \end{cases}$$

[1] 在 $\mathbb{R}^d$ 上用 $\langle\cdot,\cdot\rangle$ 表示内积, 它对保留这个常规非常有用: 在无限维背景下, 通常使用 $(\cdot,\cdot)$ 表示内积, 用 $\langle\cdot,\cdot\rangle_{H^*,H}$ 表示 H 和 H^* 之间的对偶对.

在先前的定义中使用了如下记号：

$$\operatorname*{ess\,sup}_{\Omega} = \inf\{C, |f| \leqslant C \text{ a.e. 在 } \Omega \text{ 上}\},$$

其中 a.e. 表示针对于 Lebesgue 测度. 有时把上述范数简写称 $\|\cdot\|_{L^p}$. 对于测度函数 $f:\Omega\to\mathbb{R}^d$, 这种范数很容易进行扩展, 在积分意义下, 通过定义 $\mathbb{R}^d$ 上的向量空间 ℓ^p 的范数来代替 $|f|$. 同样, 也可以使用算子 ℓ^p 范数来考虑测度 $f:\Omega\to\mathbb{R}^{d\times d}$. 在所有的这些例子中, $L^p(\Omega;X)$ 可以简记为 $L^p(\Omega)$, 其中 $X=\mathbb{R},\mathbb{R}^d,\mathbb{R}^{d\times d}$. 至于更广义的 Banach 空间值函数 f, 见定义 2.22. 本小节中的其他讨论也可以根据这个定义进行扩展.

定义 2.13 $L^p(\Omega)$ 表示所有测度函数① $f:\Omega\to\mathbb{R}$ 满足 $\|f\|_{L^p(\Omega)}<\infty$ 所构成的向量空间.

定理 2.14 (L^p 空间的基本性质)

(i) 向量空间 $L^p(\Omega)(p\in[1,\infty])$ 空间是 Banach 空间;

(ii) 如果在 $L^2(\Omega)$ 空间中定义内积

$$(u,v)_{L^2(\Omega)}=\int_\Omega u(x)v(x)\mathrm{d}x,\quad \forall u,v\in L^2(\Omega),$$

则 $L^2(\Omega)$ 空间是 Hilbert 空间;

(iii) $L^p(\Omega)$ 是可分的 ($p\in[1,\infty)$) 和自反的 ($p\in(1,\infty)$), 但是 $L^1(\Omega)$ 空间不是自反的, 并且 $L^\infty(\Omega)$ 既不是自反的也不是可分的.

设 $p\in[1,\infty]$, 通过如下式子：

$$\frac{1}{p}+\frac{1}{q}=1 \tag{2.4.1}$$

定义 $q\in[1,\infty]$, 则 Hölder 不等式成立, 即

$$\left|\int_\Omega u(x)v(x)\mathrm{d}x\right|\leqslant\|u\|_{L^p(\Omega)}\|v\|_{L^q(\Omega)},\quad \forall u\in L^p(\Omega),v\in L^q(\Omega).$$

设 $p\in[1,\infty)$, 通过 (2.4.1) 可定义 q, 则

$$(L^p(\Omega))^*=L^q(\Omega).$$

L^2 内积很容易扩展到 $\mathbb{R}^d$ 和 $\mathbb{R}^{d\times d}$ 上的函数, 在积分意义下, 通过定义 $\mathbb{R}^d$ 或 $\mathbb{R}^{d\times d}$ 上的内积可以代替逐点相乘, 在这种情况下, 使用记号 $L^2(\Omega)$.

根据定理 2.14 的最后一个部分和 $L^p(\Omega)$ 是 Banach 空间以及定义 2.6, 暗含了当 $p\in[1,\infty)$ 时, 空间 $L^p(\Omega)$ 上弱收敛的如下等价定义：

① 严格来讲, $L^p(\Omega)$ 是函数等价类的集合, 其中由等式 a.e. 定义的等价类表示 Lebesgue 测度.

定义 2.15　序列 $\{u_n\}_{n=1}^{\infty} \subset L^p(\Omega)\ (p \in [1,\infty))$ 满足如下条件:

$$\int_{\Omega} u_n(x)v(x)\mathrm{d}x \to \int_{\Omega} u(x)v(x)\mathrm{d}x, \quad \forall v \in L^q(\Omega),$$

其中 q 的定义见 (2.4.1), 则称 $\{u_n\}_{n=1}^{\infty}$ 弱收敛于 $u \in L^p(\Omega)$, 可以写成

$$u_n \rightharpoonup u \text{ 弱 } L^p(\Omega).$$

注意到当 $p \in (1,\infty)$ 时, $L^p(\Omega)$ 是自反的. 因此, 在这个空间中的每一个有界序列都有一个弱收敛子列.

为了简单起见, 用 $(\cdot,\cdot)$ 表示 $L^2(\Omega)$ 空间中的内积, 用密度参量去简化弱收敛性的验证. 例如, 为了证明在 $L^2(\Omega)$ 中 u_n 弱收敛于 u, 根据 $C(\bar{\Omega})$ 在 $L^2(\Omega)$ 中稠密, 只需要证明

$$(u_n - u, v) \to 0, \quad \forall v \in C(\bar{\Omega}).$$

鉴于极少用到弱收敛 L^∞(因为它的对偶空间不是 L^p 空间), 所以此空间中弱 $*$ 收敛用处很广 (因为其对偶空间为 L^1, 是一个可分空间).

定义 2.16　序列 $\{u_n\}_{n=1}^{\infty} \subset L^\infty(\Omega)$, 如果

$$\int_{\Omega} u_n(x)\phi(x)\mathrm{d}x \to \int_{\Omega} u(x)\phi(x)\mathrm{d}x, \quad \forall \phi \in \mathrm{L}^1(\Omega),$$

则称 $\{u_n\}_{n=1}^{\infty}$ 在 L^∞ 中弱 $*$ 收敛, 记为

$$u_n \overset{*}{\rightharpoonup} u \text{ 弱 } * \text{ 于 } L^p(\Omega) \text{ 上}.$$

因为 $L^1(\Omega)$ 是可分的, 在以往定义的基础上可以推导出 $L^\infty(\Omega)$ 中的每一个有界序列都有一个弱 $*$ 收敛的子序列.

2.4.3 Sobolev 空间

定义 2.17　设 $u, v \in L^2(\Omega)$. 如果满足

$$\int_{\Omega} u\frac{\partial \phi}{\partial x_i}\mathrm{d}x = -\int_{\Omega} v\phi\mathrm{d}x, \quad \forall \phi \in \mathrm{C}_0^\infty(\Omega),$$

则称 v 是 u 关于 x_i 的弱导数.

类似地, 可以定义高阶弱导数, 基于函数 $u \in L^2(\Omega)$ 以前的定义, 使用符号 $\partial u/\partial x_i$ 表示关于 x_i 的弱导数, 用符号 $\nabla u = \sum\limits_i \partial u/\partial x_i e_i$ 表示其梯度.

定义 2.18　Sobolev 空间 $H^1(\Omega)$ 包含所有的平方可积函数 $u: \Omega \to \mathbb{R}$, 其一阶弱导数存在且平方可积,

$$H^1(\Omega) = \{u | u, \nabla u \in L^2(\Omega)\}.$$

$H^1(\Omega)$ 是可分的 Hilbert 空间, 含内积

$$(u,v)_{H^1(\Omega)} = (u,v)_{L^2(\Omega)} + (\nabla u, \nabla v)_{L^2(\Omega)}$$

以及范数

$$\|u\|_{H^1(\Omega)} = \left(\|u\|^2_{L^2(\Omega)} + \|\nabla u\|^2_{L^2(\Omega)}\right)^{1/2}. \tag{2.4.2}$$

由于 $H^1(\Omega)$ 是一个 Hilbert 空间, 它是自反的, 因此, 每一个有界序列 $H^1(\Omega)$ 都有一个收敛子序列. 实际上, 更确切地说, $H^1(\Omega)$ 具有如下一个非常有用的性质: 它到 $L^2(\Omega)$ 的嵌入是紧的, 这意味着以下内容是成立的:

定理 2.19 (Rellich 紧性定理) $H^1(\Omega)$ 中的每一个有界序列中都存在一个在 $L^2(\Omega)$ 中强收敛的子序列.

在很多情况下, $H^1(\Omega)$ 中的元素需要考虑在区域边界 $\partial\Omega$ 上为零, 因此, 具有这种性质的函数构成 $H^1(\Omega)$ 的一个如下子集:

定义 2.20 定义 Sobolev 空间 $H_0^1(\Omega)$ 是基于 $H^1(\Omega)$ 范数而得到的 $C_0^\infty(\Omega)$ 的完备化空间.

$H_0^1(\Omega)$ 一个重要的性质就是可以根据 L^2 范数的梯度控制 L^2 范数的元素, 现在只提出这个结果. 在以后的部分和章节中, 将其扩展为 $H^1\left(\mathbb{T}^d\right)$ 周期元素的模拟.

定理 2.21 (Poincaré 不等式) 设 Ω 是 $\mathbb{R}^d$ 上的一个有界开集, 则存在一个常量 C_Ω, 它只依赖于 Ω, 对于任意的 $u \in H_0^1(\Omega)$ 满足

$$\|u\|_{L^2(\Omega)} \leqslant C_\Omega \|\nabla u\|_{L^2(\Omega)}. \tag{2.4.3}$$

定理 2.21 第一部分 (见练习第 9 题) 的直接推论, 用 $\|\nabla \cdot\|_{L^2(\Omega)}$ 范数替代 $H_0^1(\Omega)$ 中的范数可得

$$\|u\|_{H_0^1(\Omega)} = \|\nabla u\|_{L^2(\Omega)}. \tag{2.4.4}$$

用 $H^{-1}(\Omega)$ 表示对偶空间 $H_0^1(\Omega)$, 并且用 $\langle\cdot,\cdot\rangle_{H^{-1},H_0^1}$ 表示 $H^{-1}(\Omega)$ 和 $H_0^1(\Omega)$ 之间元素的一种关系, 换言之, $f \in H^{-1}(\Omega)$ 在 $v \in H_0^1(\Omega)$ 上的作用可以表示为 $\langle f,v\rangle_{H^{-1},H_0^1}$, 则 $H^{-1}(\Omega)$ 是具有如下范数:

$$\|f\|_{H^{-1}} = \sup_{v\in H_0^1\neq 0} \frac{\left|\langle f,v\rangle_{H^{-1},H_0^1}\right|}{\|v\|_{H_0^1}} = \sup_{v\in H_0^1, \|v\|_{H_0^1}\leqslant 1} \left|\langle f,v\rangle_{H^{-1},H_0^1}\right|$$

的 Banach 空间. 这样, 如下 Cauchy-Schwarz 不等式成立:

$$\left|\langle f,v\rangle_{H^{-1},H_0^1}\right| \leqslant \|f\|_{H^{-1}} \|v\|_{H_0^1}, \quad \forall f \in H^{-1}(\Omega),\ \forall v \in H_0^1(\Omega). \tag{2.4.5}$$

2.4.4 Banach 空间值空间

通过用定义在 Ω 上的另外一种测度的积分来代替 Lebesgue 积分, 使得 L^p 空间里面的函数比定义在 $\mathbb{R}^d$ 上的函数更加广泛, 从而使得 L^p 空间里的函数可以在任意 Banach 空间上取值, 下面将说明这种思想.

这里定义的空间经常在应用中出现, 特别地, 周期均匀化方法的严密分析非常需要处理两参数的函数 (一个参数变化慢, 另外一个变化非常快).

定义 2.22　设 X 是具有范数 $\|\cdot\|_X$ 的 Banach 空间, Ω 表示 $\mathbb{R}^d$ 的子集 (不一定有界), 空间 $Y := L^p(\Omega; X)\,(p \in [1,+\infty])$ 是由所有的测度函数 $u : x \in \Omega \to u(x) \in X$ 满足 $\|u(x)\|_X \in L^p(\Omega)$ 组成的.

空间 Y 具有许多好的性质, 列举如下:

定理 2.23　设 $Y = L^p(\Omega; X)$, 其中 X 和 Ω 如定义 2.22 所定义, 则

(i) Y 关于范数 $\|u\|_Y := \begin{cases} \left(\displaystyle\int_\Omega \|u(x)\|_X^p \mathrm{d}x\right)^{1/p}, & 1 \leqslant p \leqslant \infty, \\ \operatorname*{ess\,sup}_{x\in\Omega} \|u(x)\|_X, & p = \infty \end{cases}$ 是 Banach 空间;

(ii) 如果 X 是自反的且 $p \in (1,+\infty)$, 则 Y 也是自反的;

(iii) 如果 X 是可分的且 $p \in [1,+\infty)$, 则 Y 也是可分的.

如下的两个例子将会显示这种思想是如何被推广的:

例 2.24　设 X 是 Banach 空间, 则可以像定理 2.23 一样定义 Banach 空间 $Y = H^1(\Omega; X)$, 这样一个空间上的范数可表示为

$$\|u\|_Y = \left(\int_\Omega \left(\|\nabla_x u(x)\|_X^2 + \|u(x)\|_X^2\right) \mathrm{d}x\right)^{1/2}. \qquad \square$$

例 2.25　如果 $(\Omega, \mathcal{F}, \mu)$ 是一个概率空间 (见第 3 章), 类似于定理 2.23(i), 采用 Ω 上概率测度 μ 定义的 L^p 积分来定义空间 $L^p(\Omega; X)$, 则 $L^p(\Omega; X)$ 将表示一个 Banach 空间. 此空间采用的不是之前的 Lebesgue 测度 (见第 3 章). 在这个例子中, 强调 L^p 积分中的测度能够被理解, 如果当 X 空间意思明了时, 可以写成 $L^p(\Omega; \mu)$; 如果 μ 是 Lebesgue 测度, 则也可以写成 $L^p(\Omega; \mathrm{Leb})$. □

下面将通过使用连续函数记号来构建如下 Banach 空间:

定义 2.26　设 $\mathcal{I}$ 表示开或者闭集 $\Omega \in \mathbb{R}$ 上的一个开或闭的区间, 则

(i) $C(\mathcal{I}; X)$ 表示从 $\mathcal{I}$ 到 Banach 空间 X 的连续函数所构成的 Banach 空间;

(ii) $C^{p,k}(\bar{\Omega} \times \mathcal{I}, \mathbb{R})$ 表示具有 k 次和 p 次连续微分对实值函数所构成的空间.

例 2.27　设 $T > 0$, 则函数空间 $L^p((0,T); X)$, $C^{p,k}(\bar{\Omega}\times[0,T]; \mathbb{R})$ 和 $C([0,T]; X)$ 将分别用来分析第 7 章中的演化偏微分方程、抛物型偏微分方程和双曲型偏微分方程.

2.4.5 周期函数的 Sobolev 空间

设 $\mathbb{T}^d$ 表示 d 维的单位圆环, 有时称为单胞, 函数 $f:\mathbb{R}^d\to\mathbb{R}$ 满足

$$f(y+e_i)=f(y),\quad \forall y\in\mathbb{R}^d, i\in\{1,\cdots,d\},$$

则称为 1 周期函数. 这样也可以认为 f 是定义在 $\mathbb{T}^d$ 到 $\mathbb{R}$ 上的函数. 在本节中, 将研究 1 周期函数 $f:\mathbb{T}^d\to\mathbb{R}$ 的一些特性.

用 $C^\infty_{\mathrm{per}}(\mathbb{T}^d)$ 表示限制在 $\mathbb{T}^d$ 上的 1 周期光滑函数 $C^\infty(\mathbb{R}^d)$, 则在 $C^\infty_{\mathrm{per}}(\mathbb{T}^d)$ 上用 L^p 范数形成的完备化空间为 $L^p_{\mathrm{per}}(\mathbb{T}^d)$. 类似地可定义 $H^1_{\mathrm{per}}(\mathbb{T}^d)$.

Poincaré 不等式在空间 H^1_{per} 内并不成立. 然而, 如果从这个空间中移走一个常数, 用这种思想可以定义

$$H=\left\{u\in H^1_{\mathrm{per}}\left(\mathbb{T}^d\right)\middle|\int_{\mathbb{T}^d}u\mathrm{d}y=0\right\},\tag{2.4.6}$$

则存在一个常数 $C_p>0$, 使得

$$\|u\|_{L^2(\mathbb{T}^d)}\leqslant C_p\|\nabla u\|_{L^2(\mathbb{T}^d)},\quad \forall u\in H.\tag{2.4.7}$$

因此, 可以使用

$$\|u\|_H=\|\nabla_u\|_{L^2(\mathbb{T}^d)},\quad u\in H\tag{2.4.8}$$

作为 H 中的范数. H 的对偶包含 $\left(H^1_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)^*$ 中所有与常数正交的元素, 即

$$H^*=\left\{u\in\left(H^1_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)^*\middle|\langle u,1\rangle_{(H^1_{\mathrm{per}})^*,H^1_{\mathrm{per}}}=0\right\}.\tag{2.4.9}$$

空间 $L^2\left(\Omega;L^2\left(\mathbb{T}^d\right)\right):=L^2\left(\Omega\times\mathbb{T}^d\right)$ 将会在研究周期性均匀化时用到, 这是一个具有内积 $(u,v)_{L^2(\Omega\times\mathbb{T}^d)}=\int_\Omega\int_{\mathbb{T}^d}u(x,y)v(x,y)\,\mathrm{d}y\mathrm{d}x$ 的 Hilbert 空间, 它相应的范数为

$$\|u\|^2_{L^2(\Omega\times\mathbb{T}^d)}=\int_\Omega\int_{\mathbb{T}^d}|u(x,y)|^2\,\mathrm{d}y\mathrm{d}x.$$

偶尔, 将要用到所有函数的集合 $u\in L^2\left(\Omega;C_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)$, 通过定理 2.23, 定义在这个空间的范数为

$$\|u\|^2_{L^2(\Omega;C_{\mathrm{per}}(\mathbb{T}^d))}=\int_\Omega\left(\sup_{y\in\mathbb{T}^d}|u(x,y)|\right)^2\mathrm{d}x.$$

这是一个可分的 Banach 空间, 它在 $L^2\left(\Omega\times\mathbb{T}^d\right)$ 空间中稠密. 在这种情况下, $u\in L^2\left(\Omega\times\mathbb{T}^d\right)$ 在第二项中变化非常迅速, 将会出现一些有用的结论, 下面将详细给出.

定理 2.28 设 $u \in L^2\left(\Omega; C_{\text{per}}\left(\mathbb{T}^d\right)\right)$, $\varepsilon > 0$, 并且定义 $u^\varepsilon(x) = u\left(x, x/\varepsilon\right)$, 则

(i) $u^\varepsilon \in L^2(\Omega)$, 并且 $\|u^\varepsilon\|_{L^2(\Omega)} \leqslant \|u\|_{L^2(\Omega; C_{\text{per}}(\mathbb{T}^d))}$;

(ii) 当 $\varepsilon \to 0$ 时, $u^\varepsilon(x)$ 在 $L^2(\Omega)$ 中弱收敛于 $\int_{\mathbb{T}^d} u(x,y)\mathrm{d}y$;

(iii) 当 $\varepsilon \to 0$ 时, $\|u^\varepsilon\|_{L^2(\Omega)} \to \|u\|_{L^2(\Omega\times\mathbb{T}^d)}$.

定理 2.28 中的结论捕捉到了这种观念: 如果振荡是有规律的, 通过渐近的方法, 快速振荡能从慢的变量中间解耦出来. 这种思想是偏微分方程均匀化理论的核心. 另外, 周期函数的如下特性将会在后面使用到.

定理 2.29 设 $p \in [1,\infty]$, $f \in L^p_{\text{per}}(\mathbb{T}^d)$, 如在 $\mathbb{R}^d$ 上, 令

$$f^\varepsilon(x) = f\left(\frac{x}{\varepsilon}\right).$$

如果 $\varepsilon \to 0$ 且 $p < \infty$, 对 $\mathbb{R}^d$ 中任意的有界开集 Ω, 在 $L^2(\Omega)$ 中有如下弱收敛:

$$f^\varepsilon \rightharpoonup \int_{\mathbb{T}^d} f(y)\mathrm{d}y,$$

在 $L^\infty\left(\mathbb{R}^d\right)$ 上, 也有如下弱 $*$ 收敛:

$$f^\varepsilon \rightharpoonup \int_{\mathbb{T}^d} f(y)\mathrm{d}y.$$

2.5 双尺度收敛

特别适合于周期问题均匀化的一种弱收敛形式就是**双尺度收敛**. 本节将给出其定义并研究其基本特性. 再次考虑 $\mathbb{T}^d$ 上的周期函数, 同前面一样, Ω 表示 $\mathbb{R}^d$ 上的子集 (不一定有界①). 首先讨论稳态问题的双尺度收敛, 接着讨论瞬态问题.

2.5.1 稳态问题的双尺度收敛

定义 2.30 设 u^ε 是 $L^2(\Omega)$ 中的一个序列, 称 u^ε 双尺度收敛于 $u_0(x,y) \in L^2\left(\Omega \times \mathbb{T}^d\right)$, 并且可以写为 $u^\varepsilon \xrightarrow{2} u_0$, 如果对每一个测试函数 $\phi \in L^2\left(\Omega; C_{\text{per}}\left(\mathbb{T}^d\right)\right)$ 都有

$$\lim_{\varepsilon\to o} \int_\Omega u^\varepsilon(x)\phi\left(x, \frac{x}{\varepsilon}\right)\mathrm{d}x = \int_\Omega \int_{\mathbb{T}^d} u_0(x,y)\,\phi(x,y)\mathrm{d}y\mathrm{d}x. \tag{2.5.1}$$

在许多实例中, 有必要通过一个更广泛的测试函数来测试序列 u^ε. 为了这个目的, 需要充分地考虑下面引理所描述的这类函数.

① 例如, 已证明了的引理 2.34 只适用于有界域 Ω, 但是也可以扩展到 Ω 无界的情况.

引理 2.31 设 $u^\varepsilon \in L^2(\Omega)$ 是一个双尺度收敛序列, 则 (2.5.1) 对于所有的测试函数 $\phi(x,y)$ 都是收敛, 其中 $\phi(x,y)$ 是由 $\psi_1(y)\psi_2(x,y)$ 构成的, 并且 $\psi_1 \in L^\infty_{\mathrm{per}}(\mathbb{T}^d)$, $\psi_2 \in L^2(\Omega; C_{\mathrm{per}}(\mathbb{T}^d))$.

双尺度函数收敛暗含了 $L^2(\Omega)$ 上的弱收敛, 特别地有如下引理:

引理 2.32 设 u^ε 是 $L^2(\Omega)$ 上的一个序列, 它双尺度收敛于 $u_0 \in L^2(\Omega \times \mathbb{T}^d)$, 则在 $L^2(\Omega)$ 上有弱收敛

$$u^\varepsilon \rightharpoonup \bar{u}_0, \tag{2.5.2}$$

其中

$$\bar{u}_0(x) = \int_{\mathbb{T}^d} u_0(x,y)\mathrm{d}y.$$

证明 任取一个测试函数 $\phi(x) \in L^2(\Omega)$, 独立于 y, 用式 (2.5.1) 推导可得

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_\Omega u^\varepsilon(x)\phi(x)\mathrm{d}x &= \int_\Omega\int_{\mathbb{T}^d} u_0(x,y)\,\phi(x)\mathrm{d}y\mathrm{d}x\\ &= \int_\Omega\left(\int_{\mathbb{T}^d} u_0(x,y)\,\mathrm{d}y\right)\phi(x)\mathrm{d}x\\ &= (\bar{u}_0,\phi)_{L^2(\Omega)}.\end{aligned}$$

这对于每一个 $\phi \in L^2(\Omega)$ 都成立, 因此, u^ε 在 $L^2(\Omega)$ 中弱收敛于 $\bar{u}_0$. □

根据引理 2.32 可直接得到如下推论:

推论 2.33 设 u^ε 是 $L^2(\Omega)$ 上的一个序列, 并且双尺度收敛于 $u_0 \in L^2(\Omega)$, 即双尺度极限依赖于 y, 那么弱 L^2 极限和双尺度极限是一致的.

双尺度收敛对于第 12~14 章的多尺度展开式形式的发展是一种很好的研究工具, 在这些章节中, 存在一个依赖于 "快变量" 的周期性, 并且在第 19~21 章中, 对这些展开式进行了严密的分析. 如下结论说明了双尺度收敛在本章中的作用:

引理 2.34 考虑函数 $u^\varepsilon \in C(\bar{\Omega})$ 的如下形式:

$$u^\varepsilon(x) = u_0\left(x,\frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x,\frac{x}{\varepsilon}\right),$$

其中 $u_j \in C(\bar{\Omega}; C_{\mathrm{per}}(\mathbb{T}^d))\,(j=0,1)$, Ω 是 $\mathbb{R}^d$ 上一个有界开集, 则有 $u^\varepsilon \xrightarrow{2} u_0$.

证明 设 $\phi \in L^2(\Omega; C_{\mathrm{per}}(\mathbb{T}^d))$, 并定义 $f_j(x,y) = u_j(x,y)\phi(x,y)\,(j=0,1)$. 使用记号 $f^\varepsilon(x) = f(x,x/\varepsilon)$, 显然有

$$\int_\Omega u^\varepsilon(x)\phi\left(x,\frac{x}{\varepsilon}\right)\mathrm{d}x = \int_\Omega f_0^\varepsilon(x)\mathrm{d}x + \varepsilon\int_\Omega f_1^\varepsilon(x)\mathrm{d}x. \tag{2.5.3}$$

当 $j=0,1$ 时有 $f_j^\varepsilon \in L^2(\Omega; C_{\mathrm{per}}(\mathbb{T}^d))$, 从而由定理 2.28 可知 f_0^ε 在 $L^2(\Omega)$ 中

弱收敛于 $\mathbb{T}^d$ 的均值 $\bar{f}_0(x) := \int_{\mathbb{T}^d} f_0(x,y)\,\mathrm{d}y$. 这个弱收敛对于所有的 $\psi \in L^2(\Omega)$ 都有

$$(\psi, f_0^\varepsilon) \to (\psi, \bar{f}_0)$$

令 $\psi = 1$, 则有

$$\begin{aligned}\int_\Omega f_0^\varepsilon(x)\mathrm{d}x &\to \int_\Omega \int_{\mathbb{T}^d} f_0(x,y)\,\mathrm{d}y\mathrm{d}x\\ &= \int_\Omega \int_{\mathbb{T}^d} u_0(x,y)\,\phi(x,y)\,\mathrm{d}y\mathrm{d}x.\end{aligned}$$

考虑 (2.5.3) 右边的第二个积分式, 由定理 2.28 可知 f_1^ε 在 $L^2(\Omega)$ 中弱收敛, 因此, 由定理 2.7 可知其在 $L^2(\Omega)$ 上有界, 那么由 Ω 的收敛性以及 Cauchy-Schwarz 不等式可得

$$\varepsilon\left|\int_\Omega f_1^\varepsilon(x)\mathrm{d}x\right| \leqslant \varepsilon C\|f_1^\varepsilon\|_{L^2(\Omega)} \to 0.$$

对于所有的 $\phi \in L^2(\Omega; C_{\mathrm{per}}(\mathbb{T}^d))$, 由 (2.5.3) 中的两个计算式可得

$$\begin{aligned}\int_\Omega u^\varepsilon(x)\phi\left(x,\frac{x}{\varepsilon}\right)\mathrm{d}x &= \int_\Omega (f_0^\varepsilon(x) + \varepsilon f_1^\varepsilon(x))\mathrm{d}x\\ &\to \int_\Omega \int_{\mathbb{T}^d} u_0(x,y)\phi(x,y)\,\mathrm{d}y\mathrm{d}x.\end{aligned}$$

因此, u^ε 双尺度收敛于 u_0. □

希望找到这样一个判别标准: 能判别 $L^2(\Omega)$ 中的序列是否是双尺度收敛. 以下介绍的紧性结论提供了这个判别标准:

定理 2.35 设 u^ε 是 $L^2(\Omega)$ 中的一个有界序列, 则存在一个由 u^ε 表示的子序列和一个函数 $u_0 \in L^2(\Omega \times \mathbb{T}^d)$, 使得 u^ε 双尺度收敛于 u_0.

这个结论关注的是 $L^2(\Omega)$ 上的有界序列, 这个序列的双尺度极限是 $L^2(\Omega \times \mathbb{T}^d)$ 中的一个元素, 并且显示地依赖于 y. 很自然地会有人问及, 当序列采用一个更强的有界范数时, 是否可以从中获得关于双尺度极限的更多信息. 回顾由 (2.4.6) 所定义的零均值函数 $H^1(\Omega)$ 的子空间 H, 可得如下结论:

定理 2.36 (i) 设 u^ε 是 $H^1(\Omega)$ 中的一个有界序列, 考虑在 $H^1(\Omega)$ 上的弱极限为 u 的子序列, 则随着这个子序列, u^ε 双尺度收敛于弱 $H^1(\Omega)$ 极限 u. 此外, 存在一个函数 $u_1 \in L^2(\Omega; H)$, 沿着子序列, 使得 ∇u^ε 双尺度收敛于 $\nabla_x u + \nabla_y u_1$.

(ii) 设 u^ε 和 $\varepsilon\nabla u^\varepsilon$ 均为 $L^2(\Omega)$ 中的有界序列, 则沿着一个子序列存在函数 $u_0 \in L^2(\Omega; H)$, 使得 u^ε 和 $\varepsilon\nabla u^\varepsilon$ 分别双尺度收敛于 $u_0(x,y)$ 和 $\nabla_y u_0$.

(iii) 设 u^ε 和 $\varepsilon^\gamma\nabla u^\varepsilon\,(\gamma \in (0,1))$ 均为 $L^2(\Omega)$ 中的有界序列, 则 u^ε 的双尺度极限依赖于 y, 并且存在一个函数 $u_0 \in L^2(\Omega)$, 使得 $u^\varepsilon \xrightarrow{2} u_0$.

定理 2.36 非常有用, 因为如果一个函数有如下形式:

$$u^{\varepsilon}(x)=u(x)+\varepsilon u_1(x,x/\varepsilon)+O\left(\varepsilon^2\right),$$

则对 $O(\varepsilon^2)$ 进行合理的假设, 希望得到对于任意 ε, $u(x)$ 都近似于 u^{ε}, $\nabla u(x)+\nabla_y u_1(x,y)\big|_{y=x/\varepsilon}$ 近似于 u^{ε} 的微分. (i) 就是用来精确处理这种情况的.

(i) 标注双尺度收敛暗含了序列 u^{ε} 在 $L^2(\Omega)$ 上弱收敛于 u, 这是因为极限 $u(x)$ 仅取决于 x 而不是 y(推论 2.33). 事实上, 由 Rellich 紧性定理可知, 采用一个更为严格的子序列, u^{ε} 在 L^2 上是强收敛的. 如果这个极限是唯一的, 则在 $H^1(\Omega)$ 中, 整个序列 u^{ε} 弱收敛于 u, 并且在 $L^2(\Omega)$ 上强收敛. 从 (iii) 中可以看到, 通过控制 u^{ε} 的梯度, 可以推导出 u 的双尺度极限独立于 y. 然而, 这个性质对于梯度的双尺度极限却不成立.

2.5.2 时变问题的双尺度收敛

我们认为研究进化偏微分方程均匀化问题时, 有必要将函数序列的时变性考虑在内, 并修改双尺度的概念, 现给出相关的定义和定理. 正如 2.5.1 小节中, 设 Ω 是 $\mathbb{R}^d$ 上的一个子集 (不一定有界), $\mathbb{T}^d$ 表示单位体, 并设 $\Omega_T=\Omega\times(0,T)$, 用 (x,y,t) 表示 $\Omega\times\mathbb{T}^d\times(0,T)=\Omega_T\times\mathbb{T}^d$ 上的一个点, 回顾 H 由 (2.4.6) 给出.

定义 2.37 设序列 $u^{\varepsilon}\in L^2\left(\Omega_T\right)$, $u_0(x,y,t)\in L^2\left(\Omega_T\times\mathbb{T}^d\right)$, 如果对于每一个函 $\phi(x,y,t)\in L^2\left(\Omega_T;C_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)$ 都有

$$\lim_{\varepsilon\to 0}\int_0^T\int_{\Omega}u^{\varepsilon}(x,t)\phi\left(x,\frac{x}{\varepsilon},t\right)\mathrm{d}x\mathrm{d}t=\int_0^T\int_{\Omega}\int_{\mathbb{T}^d}u_0(x,y,t)\phi(x,y,t)\mathrm{d}y\mathrm{d}x\mathrm{d}t,\quad(2.5.4)$$

则称序列 u^{ε} 双尺度收敛于 $u_0(x,y,t)$, 记为 $u^{\varepsilon}\xrightarrow{2}u_0$.

注意到在先前的定义中, t 仅是作为一个参数被计入的. 然而在时变性问题中, 需要在 (2.5.4) 中使用比 $L^2\left(\Omega_T;C_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)$ 中更为广泛的测试函数. 例如, 可以将形式为 $\psi_1(y)\psi_2(x,y,t)\psi\in L^{\infty}_{\mathrm{per}}\left(\mathbb{T}^d\right)$ 和 $\phi\in L^2\left(\Omega_T;C_{\mathrm{per}}\left(\mathbb{T}^d\right)\right)$ 的测试函数用于式 (2.5.4).

最基本的双尺度收敛的紧性定理, 即定理 2.35 仍然是有效的.

定理 2.38 设 u^{ε} 是 $L^2\left(\Omega_T\right)$ 中的一个有界序列, 那么存在一个子序列 (仍然用 u^{ε} 表示) 以及一个函数 $u_0\in L^2\left(\Omega_T\times\mathbb{T}^d\right)$, 使得 u^{ε} 双尺度收敛于 u_0, 而且 u^{ε} 通过单胞

$$u^{\varepsilon}\rightharpoonup\int_{\mathbb{T}^d}u_0(\cdot,y,\cdot)\mathrm{d}y$$

在 $L^2(\Omega_T)$ 上弱收敛于其双尺度极限的均值. 在稳态情况下, 已经知道, 更好空间上的有界性能够提供更多关于双尺度极限的信息. 以下定理类似于定理 2.36:

定理 2.39 (i) 设 u^ε 是 $L^2((0,T),H^1(\Omega))$ 中的有界序列, 考虑 $L^2((0,T), H^1(\Omega))$ 上的弱极限为 u 的子序列, 则沿着这个子序列, 在 $L^2(\Omega_T;H)$ 上, u^ε 双尺度弱收敛于 u. 此外, 存在一个函数 $u_1 \in L^2(\Omega_T;H)$, 沿着子序列, 使得 ∇u^ε 双尺度收敛于 $\nabla_x u + \nabla_y u_1$.

(ii) 设 u^ε 和 $\varepsilon\nabla u^\varepsilon$ 均为 $L^2(\Omega_T)$ 中的有界序列, 那么沿着子序列, 存在函数 $u_0 \in L^2(\Omega_T;H)$, 使得 u^ε 和 $\varepsilon\nabla u^\varepsilon$ 分别双尺度收敛于 u_0 和 $\nabla_y u_0$.

(iii) 设 u^ε 和 $\varepsilon^\gamma\nabla u^\varepsilon\ (\gamma\in(0,1))$ 均为 $L^2(\Omega_T)$ 中的有界序列, 则 u^ε 的双尺度极限依赖于 y, 并且存在一个函数 $u_0 \in L^2(\Omega_T)$, 使得

$$u^\varepsilon \stackrel{2}{\rightharpoonup} u_0.$$

这两个定理的证明与前面章节的证明基本上是相同的.

2.6 Hilbert 空间中的方程

在适当的 Hilbert 空间中, 许多偏微分方程都可以用等式的形式非常便利地表示出来. 这对本书中要研究的线性椭圆型及线性抛物型偏微分方程是一个很有用的观点, 从而对抽象这些问题也是非常有用的. 这里总结一下这个理论, 它包括两个重要定理: Lax-Milgram 定理和 Fredholm 替换, 其中 Lax-Milgram 定理为 Hilbert 空间中的线性方程提供了解的存在性和唯一性.

2.6.1 Lax-Milgram 定理

设 H 是内积为 $(\cdot,\cdot)$ 的 Hilbert 空间①, $A: H\to H^*$ 是一个线性算子. 又设 $f\in H^*$ 并且 $\langle\cdot,\cdot\rangle_{H^*,H}$ 表示 H 和 H^* 之间的对偶, 研究方程

$$Au = f. \tag{2.6.1}$$

方程 (2.6.1) 的弱形式是

$$(Au, v) = \langle f, v\rangle_{H^*,H}, \quad \forall v\in H. \tag{2.6.2}$$

A 的线性暗含了方程 $a[u,v]=(Au,v)$ 的左边项可以定义一个双线性形式 $a: H\times H\to\mathbb{R}$.

方程 (2.6.2) 的解的存在性和唯一性可以由以下定理证明:

定理 2.40 (Lax-Milgram) 设 H 是一个具有范数 $\|\cdot\|$ 以及内积 $(\cdot,\cdot)$ 的 Hilbert 空间, 用 $\langle\cdot,\cdot\rangle_{H^*,H}$ 表示 H^* 和 H 之间的对偶对, 设 $a: H\times H\to\mathbb{R}$ 是一个双线性映射, 满足如下条件:

---

① 注意到这里讨论的是普通的 Hilbert 空间, 而不是 (2.4.6) 给出的特定 H.

(i)(椭圆性) 存在一个常数 $\alpha > 0$, 使得

$$a[u,u] \geqslant \alpha\|u\|^2, \quad \forall u \in H;$$

(ii)(连续性) 存在一个常数 $\beta > 0$, 使得

$$a[u,v] \leqslant \beta\|u\|\|v\|, \quad \forall u,v \in H.$$

另设 $f: H \to \mathbb{R}$ 是 H 上的一个有界线性范函, 则存在唯一元素 $u \in H$, 对所有的 $v \in H$, 使得

$$a[u,v] = \langle f,v\rangle_{H^*,H}\,.$$

定理 2.40 是非常有用的, 它是椭圆型偏微分方程弱解理论的中心, 它将在第 7 章中被推广, 并且将在第 19 章中用于证明椭圆型偏微分方程的均匀化理论.

2.6.2 Fredholm 性质

假定 $A: H \to H$ 是一个线性算子, 考虑 (2.6.1) 中的 $f \in H$, 可以通过恒等式

$$(Au,v) = (u, A^*v), \quad \forall u,v \in H.$$

用 $A^*: H \to H$ 表示 A 的伴随算子. 设 $v \in H$ 为 A^* 中零空间 $\mathcal{N}(A^*)$ 的任一元素, 其中

$$\mathcal{N}(A^*) = \{v \in H | A^*v = 0\}.$$

(2.6.1) 暗含

$$(f,v) = 0, \quad \forall v \in \mathcal{N}(A^*).$$

因此, (2.6.2) 的解的存在性需要一个必要条件, 即方程右边的等式与 A 的伴随算子的零空间正交 (注意到如果 $\mathcal{N}(A^*)$ 是非空的, 则其椭圆性失效, 这样 Lax-Milgram 定理不适用).

形式参量适用于有限维, 但当 A 是恒等式 $A = I - K$(K 是紧密的) 的一个紧扰动时, 也可以适用于无限维, 从而以下的定义非常重要:

定义 2.41 如果有界算子 $K: H \to H$ 从有界集映射到紧性闭集 $\overline{K(M)}$, 即 $\overline{K(M)}$ 在 H 中对所有的有界集 $M \subset H$ 是紧的, 则称算子 K 是紧的. 它的等价定义是: 当且仅当对于每一个有界序列 $\{u_n\}_{n=1}^{\infty} \in H$, $\{Ku_n\}_{n=1}^{\infty}$ 在 H 上都有一个强收敛子列.

下面将根据 $A = I - K$ 来研究方程 (2.6.1).

定理 2.42 (Fredholm 变换) 设 H 是 Hilbert 空间, $K: H \to H$ 是一个紧算子, 则以下变换成立:

(i) 对于每个 $f, F \in H$, 如下方程:

$$(I-K)u=f, \tag{2.6.3a}$$

$$(I-K^*)U=F \tag{2.6.3b}$$

都有唯一解; 或者,

(ii) 齐次方程

$$(I-K)V_0=0, \quad (I-K^*)v_0=0$$

有相同数目的非零解, 其数目满足

$$1 \leqslant \dim(\mathcal{N}(I-K))=\dim(\mathcal{N}(I-K^*))<\infty.$$

在此情况下, 方程 (2.6.3a) 和 (2.6.3b) 有一个解的充分必要条件分别为

$$(f, v_0)=0, \quad \forall v_0 \in \mathcal{N}(I-K^*)$$

和

$$(F, V_0)=0, \quad \forall V_0 \in \mathcal{N}(I-K).$$

定理 2.42 对椭圆型边值问题的研究更为有用. 它将在第 7 章得到推广, 对它的使用也将贯穿本书的第二部分.

2.7 讨论和参考

在本书中, 采用的是无下标符号, 这种符号经常在连续介质力学的教科书中采用. 例如, 见文献 [128] 第 1 章和第 2 章以及其中的参考文献. 无下标符号使得计算公式的推导和结论的描述更加清晰. 然而, 为了验证很多的逐行计算, 读者也不用担心会拘泥于这些索引. 在数学中, 无下标的一行数学公式能够隐藏采用索引记号的多行计算.

本章中提出的大多数其他相关知识也是非常规范的, 它可以在一些泛函分析和偏微分方程的书中找到. 关于 Banach 空间、Hilbert 空间以及 L^p 空间的信息, 见文献 [53, 179, 189, 274, 337].

由于 $L^1(\Omega)$ 不是自反的, 因此, $L^1(\Omega)$ 中的有界序列不一定要有任何弱收敛子序列. 很自然地要问, 在什么情况下, 可以从属于 $L^1(\Omega)$ 空间的序列中提取一个弱收敛子序列, 这个问题可以通过 **Dunford-Pettis 定理**(除了有界还需等度可积) 找到答案, 详细参见文献 [87]. $L^1(\Omega)$ 的非自反性就是说这个空间不能表征 Hilbert 空间的对偶, 因此, 弱 $*$ 收敛对这个空间来说并不是一个有用的概念, 然而对 Ω 的

Radon 测度空间 $(C_b(\Omega))^* =: M(\Omega)$ [①]却是一个非常重要的概念. 事实上, $L^1(\Omega)$ 上的有界序列在 $M(\Omega)$ 上是弱 * 紧性的: 可以提取一个弱收敛子序列, 这个子序列收敛于 $M(\Omega)$ 中的元素. 概率学家提到的 $M(\Omega) = (C_b(\Omega))^*$ 中的弱 * 收敛就是**概率测度的弱收敛**. 这是概率理论中极限定理最有用 (最本质) 的概念, 也是 3.5 节的主题. 感兴趣的读者也可以查阅文献 [37, 156].

可以用傅里叶级数定义周期函数的 Sobolev 空间. 例如, 空间 $H^1_{\text{per}}(\mathbb{T}^d)$ 可以定义为

$$H^1_{\text{per}}\left(\mathbb{T}^d\right) = \left\{ u : u = \sum_{k\in\mathbb{Z}^d} u_k \mathrm{e}^{2\pi\mathrm{i}k\cdot x}, u_k = u_{-k}, \sum_{k\in\mathbb{Z}^d} |k|^2 |u_k|^2 < \infty \right\}.$$

类似地有

$$H = \left\{ u \in H^1_{\text{per}}\left(\mathbb{T}^d\right) : u_0 = 0 \right\}.$$

文献 [274, 315] 对周期函数的 Sobolev 空间进行了讨论; Sobolev 空间的详尽处理方法可以在文献 [4] 中找到; 偏微分方程理论的很多应用可见文献 [98, 第 5 章].

双尺度收敛的概念在文献 [233, 234] 中由 Nguetseng 给出了介绍, 之后, 在文献 [5, 6] 中由 Allaire 得以推广和发展. 2.5 节中的大多数结论源于文献 [6]. 关于双尺度收敛, 目前最新的评论文章可见文献 [197, 341].

很自然地提出一个问题: 适用于方程 (2.5.1) 的测试函数组成的最大集合是什么? 这个问题关系到容许测试函数集合的概念, 它可定义为 $\phi \in L^2(\Omega \times \mathbb{T}^d)$ 所构成的集合, 其中

$$\lim_{\varepsilon\to 0} \int_\Omega \left| \phi\left(x, \frac{x}{\varepsilon}\right) \right|^2 \mathrm{d}x = \int_\Omega \int_{\mathbb{T}^d} |\phi(x,y)|^2 \mathrm{d}y\mathrm{d}x. \tag{2.7.1}$$

由定理 2.28 可知, 任一 $\phi \in L^2\left(\Omega; C_{\text{per}}\left(\mathbb{T}^d\right)\right)$ 是容许函数. 在引理 2.31 中提到的函数也是容许函数. 然而, 更重要的是, 容许测试函数集合是 $L^2\left(\Omega \times \mathbb{T}^d\right)$ 的一个适合子集, 这并不满足 (2.7.1) 的元素 —— 见文献 [6]. 无论是变量 x 还是 y 都需要一定的正则性. 方程 (2.5.1) 中测试函数集的特性在 [142, 定理 2.3] 中给出了介绍.

针对 (2.7.1), 如果双尺度收敛序列 $u(x, x/\varepsilon)$ 同时也是一个容许测试函数, 此外, 双尺度极限 $u_0 \in L^2(\Omega; C_{\text{per}}(\Omega))$, 则 u^ε 强双尺度收敛于 u_0, 即

$$\lim_{\varepsilon\to 0} \left\| u^\varepsilon(\cdot) - u_0\left(\cdot, \frac{\cdot}{\varepsilon}\right) \right\| = 0.$$

然而, 注意到双尺度极限并不具有更为广泛的正则性. 实际上, $L^2\left(\Omega \times \mathbb{T}^d\right)$ 中的每一个函数 u_0 都可以作为 $L^2(\Omega)$ 中某一序列的双尺度极限 [6, 定理 1.13]. 对强双尺度收敛和弱双尺度收敛的相关性讨论可见文献 [341].

① 这个空间比 $L^1(\Omega)$ 更大, 它是 $M(\Omega)$ 的一个适当子集.

2.5.2 小节考虑了在时间上不产生振荡的函数序列, 通过进一步延伸双尺度收敛的概念, 得到了处理如下形式的序列:

$$u^\varepsilon(x,t) = u\left(x, \frac{x}{\varepsilon}, t, \frac{t}{\varepsilon^p}\right),$$

其中 $p > 0$, $u(x,y,t,\tau)$ 是 y 和 τ 上的周期函数, 见文献 [142]. 在空间和时间上, 研究具有快速振荡系数的抛物型偏微分方程的均匀化理论时, 这个延伸是非常有用的; 双尺度收敛的概念得到了进一步延伸, 这也导致了能够处理两种以上尺度的函数 (详见文献 [8]), 如

$$u^\varepsilon = u\left(x, \frac{x}{\varepsilon}, \frac{x}{\varepsilon^2}, \cdots\right).$$

非周期振荡系数问题的类似概念也得到了发展, 见文献 [213].

选择适当的测试函数来研究偏微分方程的渐近性问题具有标准的技巧, 如可参见文献 [95]. 在概率背景下, 由 Kurtz 在文献 [181] 中介绍了一种类似的技巧来处理双尺度收敛理论. 这一种技巧方法也在文献 [96, 97] 中被 Evans 采用以处理扰动测试函数问题.

文献 [98] 中涉及了 Lax-Milgram 定理, 而 Fredholm 变换对本书第二部分的发展占有绝对的中心位置. 在这种背景下, 这种性质在赋范空间中都成立, 而不要求空间的完备性以及内积结构, 可见文献 [179, 8.7 节].

2.8 练 习

1. 说明对于任意标量场 ϕ 都有

$$\Delta\phi = I : \nabla\nabla\phi.$$

2. 设 $\{X, \|\cdot\|\}$ 是一个 Banach 空间, 说明每一个强收敛序列都是弱收敛的.
3. 设 $\{X, \|\cdot\|\}$ 是一个有限维 Banach 空间, 说明每一个弱收敛序列都是强收敛的.
4. 设 $\Omega = (0,1) \subset \mathbb{R}$, 定义

$$u(x) = \begin{cases} x, & 0 \leqslant x \leqslant \dfrac{1}{2}, \\ 1-x, & \dfrac{1}{2} < x \leqslant 1. \end{cases}$$

说明 $u(x)$ 的弱导数是

$$\frac{\mathrm{d}u(x)}{\mathrm{d}x} = \begin{cases} 1, & 0 \leqslant x \leqslant \dfrac{1}{2}, \\ -1, & \dfrac{1}{2} < x \leqslant 1, \end{cases}$$

并说明这个函数在古典意义下是否为可微的.

5. 考虑第 4 题中的函数 u, 说明 $u \in H_0^1(\Omega)$.

6. 回顾 (2.4.6) 中 H 的定义. 使用傅里叶级数证明 Poincaré不等式在 H 上成立,

$$\|u\|_{L^2} \leqslant \frac{1}{2\pi}\|\nabla u\|_{L^2}, \quad \forall u \in H.$$

7. 设 $\Omega = (0,1) \subset \mathbb{R}$, 并定义 $u_\alpha(x) = |x|^\alpha$, 则当 α 为何值时, $u_\alpha \in H^1(\Omega)$?

8. 设 $L > 0$, 对于函数 $f \in C^\infty(0,L) \bigcap H_0^1(0,L)$, 证明定理 2.21 中的 Poincaré 不等式. 估计 Poincaré 常数 C_L 的最优值, 说明

$$\lim_{L\to\infty} C_L = \infty,$$

并解释这个结论.

9. 证明范数 (2.4.4) 等价于范数 (2.4.2).

10. 考虑函数 $u^\varepsilon \in L^2(\Omega)$ 允许如下双尺度展开式:

$$u^\varepsilon(x) = \sum_{j=0}^{N} \varepsilon^j u_j\left(x, \frac{x}{\varepsilon}\right),$$

其中 $u_j \in C(\Omega; C_{\text{per}}(Y))\,(j = 0,1,\cdots,N)$, 并且 Ω 是 $\mathbb{R}^d$ 上的有界域. 说明 $u^\varepsilon \stackrel{2}{\rightharpoonup} u_0$ (这是定理 2.34 的一个推广).

11. 考虑线性方程 (1.3.5) 和 (1.3.8) 的系统. 设 $\left\|(\mathcal{L}^\varepsilon)^{-1}\right\|$ 有界且独立于 ε, 说明在这两种情况下都有

$$\|u^\varepsilon - u_0\| \leqslant C\varepsilon.$$

第3章　概率论和随机过程

3.1　格　局

本章将给出一些有关概率论和随机过程的基本定义和结论; 还将定义 Wiener 过程 (Brownian 运动) 并发展随机积分中的 Itô理论; 总结鞅的一些特性, 并把它运用到 Itô积分中去. 本书在研究 Markov 问题中维数减少时, 经常会用到概率测度的弱收敛, 在后面的章节中, 研究随机过程的极限定理时需要用到 Markov 链和随机微分方程的平均化和均匀化, 它们与度量空间中概率测度的弱收敛性定理紧密联系, 因此, 将讨论不同种类的概率收敛形式和弱收敛形式. 就像第 2 章中说的那样, 建议大家不要一次就把书读完, 而是把本章作为以后章节学习的参考.

3.2 节和 3.3 节分别介绍了概率论和随机过程中的一些思想; 3.4 节讨论了鞅和随机积分; 3.5 节发展了弱收敛, 3.6 节给出了理论扩展和本章用到的一些参考文献.

3.2　概率论、期望和条件期望

如果集合 Ω 的子集包含 Ω, 并且其中的元素可数, 和运算都是封闭的, 则称此子集为 σ 代数. 若由 σ 代数所组成的子集满足上述公理性, 则构成子 σ 代数.

在测度空间 $(\Omega,\mathcal{F})$ 中, Ω 表示集合, $\mathcal{F}$ 是 Ω 子集的一个 σ 代数. 设 $(\Omega,\mathcal{F})$ 和 $(\mathrm{E},\mathcal{G})$ 是两个测度空间, 函数 $X:\Omega\mapsto E$ 满足如下事件:

$$\{\omega\in\Omega : X(\omega)\in A\} =: \{X\in A\}$$

并且对于每一个 $A\in\mathcal{G}$ 属于 $\mathcal{F}$, 则称 X 是可测函数或随机变量.

设 $(\Omega,\mathcal{F})$ 是一个测度空间, 如果满足 $\mu(\varnothing)=1, \mu(\Omega)=1$, 并且对于所有不相交集合所构成的序列 $\{A_k\}_{k=1}^{\infty}\in\mathcal{F}$, 满足 $\mu\left(\bigcup\limits_{k=1}^{\infty}A_k\right)=\sum\limits_{k=1}^{\infty}\mu(A_k)$, 则称函数 $\mu:\mathcal{F}\mapsto[0,1]$ 是一个概率测度, 相应地, 称 $(\Omega,\mathcal{F},\mu)$ 为概率空间. 设 X 是定义在 $(\Omega,\mathcal{F},\mu)$ 到 $(E,\mathcal{G})$ 中一个测度函数 (随机变量), 则可根据测度 μ 定义如下期望①:

$$\mathbb{E}X=\int_{\Omega}X(\omega)\mathrm{d}\mu(\omega).$$

① 通过一个极限来定义关于测度 μ 的积分, 该积分始于一个阶梯函数的定义.

更一般地, 设 $f: E \mapsto \mathbb{R}$ 表示 $\mathcal{G}$ 测度, 则

$$\mathbb{E} f(X)=\int_{\Omega} f(X(\omega)) \mathrm{d} \mu(\omega).$$

设 U 是拓扑空间, 用 $\mathcal{B}(U)$ 表示 U 的 Borel-σ 代数: 最小的 σ 代数包含 U 中的所有开集. 定义从 $(\Omega, \mathcal{F}, \mu)$ 到 $(E, \mathcal{B}(E))$ 的每个随机变量所诱导出来的概率测度 E,

$$\mu_X(B)=\mathbb{P} X^{-1}(B)=\mu(\omega \in \Omega ; X(\omega) \in B), \quad B \in \mathcal{B}(E),$$

称测度 μ_X 为 X 分布 (有时称为 X 分布规律).

例 3.1 令 $\mathcal{I}$ 表示正整数集, 如果向量 $\rho_0=\{\rho_{0,i} | i \in \mathcal{I}\}$ 含有非负通道, 并且其总质量等于 1, 即 $\sum_{i \in \mathcal{I}} \rho_{0,i}=1$, 则该向量确定一个定义在 $\mathcal{I}$ 上的测度 ($\mathcal{I}$ 值随机变量的分布). □

可以使用随机变量的分布计算期望值和概率,

$$\mathbb{E} f(X)=\int_E f(x) \mathrm{d} \mu_X(x)$$

和

$$\mathbb{P}(X \in G)=\int_G \mathrm{d} \mu_X(x), \quad G \in \mathcal{B}(E).$$

在本书中, 独立同分布随机变量的集合可以写成 i.i.d.

当 $E=\mathbb{R}^d$ 且测度 μ_X 关于 Lebesgue 测度有 Radon-Nikodym 导数, 则可记

$$\mathrm{d} \mu_X(x)=\rho(x) \mathrm{d} x.$$

其中 $\rho(x)$ 表示概率密度函数 (pdf), 或者是关于 X 的 Lebesgue 测度密度, 而且当 $E=\mathbb{R}^d$, 则用 $L^p\left(\Omega ; \mathbb{R}^d\right)$ 或 $L^p(\Omega ; \mu)$, 或简化为 $L^p(\mu)$, 可在 Ω 上用如下范数定义测度函数的 Banach 空间:

$$\|X\|_{L^p}=\left(\mathbb{E}|X|^p\right)^{1/p}.$$

例 3.2 (i) 考虑随机变量 $X: \Omega \mapsto \mathbb{R}$, 其概率密度函数为

$$\gamma_{\sigma, m}(x):=\left(2 \pi \sigma^2\right)^{-1/2} \exp \left(-\frac{(x-m)^2}{2 \sigma^2}\right).$$

此处, 称 X 为 Gaussian 或正态随机变量, 其均值为

$$\mathbb{E} X=\int_{\mathbb{R}} x \gamma_{\sigma, m}(x) \mathrm{d} x=: m,$$

并且方差为

$$\mathbb{E}(X-m)^2=\int_{\mathbb{R}}(x-m)^2\gamma_{\sigma,m}(x)\mathrm{d}x=:\sigma^2.$$

既然平均和方差完全是针对定义在 $\mathbb{R}$ 上的 Gaussian 随机变量而言的, 因此, Gaussian 可以表示为 $\mathcal{N}(m,\sigma^2)$, $\mathcal{N}(0,1)$ 是**标准正态随机变量**.

(ii) 设 $m\in\mathbb{R}^d$ 且 $\varSigma\in\mathbb{R}^{d\times d}$ 是对称正定的. 随机变量 $X:\varOmega\mapsto\mathbb{R}^d$ 具有概率密度函数

$$\gamma_{\varSigma,m}(x):=\left((2\pi)^d\det\varSigma\right)^{-1/2}\exp\left(-\frac{1}{2}\left\langle\varSigma^{-1}(x-m),(x-m)\right\rangle\right),$$

称该随机变量为**多变量**Gaussian 或**正态**随机变量, 其均值为

$$\mathbb{E}X=m \tag{3.2.1}$$

相应的协方差阵为

$$\mathbb{E}((X-m)\otimes(X-m))=\varSigma. \tag{3.2.2}$$

既然均值和方差完全是针对于定义在 $\mathbb{R}^d$ 上的 Gaussian 随机变量而言的, 因此, 通常 Gaussian 可以表示为 $\mathcal{N}(m,\varSigma)$.

(iii) 有时需要用到**退化的**Gaussian 例子, 此处 $\varSigma$ 是对称半正定的. $\varSigma$ 相对于零特征值的特征空间决定坐标方向, 在此坐标中, 概率测度是非随机的 —— 一种 **Dirac 质量**. □

例 3.3 带 $\lambda>0$ 的指数随机变量 $T:\varOmega\to\mathbb{R}^+$ 满足

$$\mathbb{P}(T>t)=\mathrm{e}^{-\lambda t},\quad\forall t\geqslant 0,$$

可记为 $T\sim\exp(\lambda)$. 相应的概率密度函数为

$$f_T(t)=\begin{cases}\lambda\mathrm{e}^{-\lambda t}, & t\geqslant 0,\\ 0, & t<0.\end{cases} \tag{3.2.3}$$

注意到

$$\mathbb{E}T=\int_{-\infty}^{\infty}tf_T(t)\mathrm{d}t=\frac{1}{\lambda}\int_0^{\infty}(\lambda t)\,\mathrm{e}^{-\lambda t}\mathrm{d}(\lambda t)=\frac{1}{\lambda}.$$

如果时间 $\tau_n=t_{n+1}-t_n$ 是关于 $\tau_0\sim\exp(\lambda)$ 的独立同分布随机变量, 令 $t_0=0$, 则有

$$t_n=\sum_{k=0}^{n-1}\tau_k,$$

并且可能会有 (见练习 3.3):

$$\mathbb{P}(0 \leqslant t_k \leqslant t < t_{k+1}) = \frac{\mathrm{e}^{-\lambda t}(\lambda t)^k}{k!}. \tag{3.2.4}$$

因此, 一个具有平均 $1/\lambda$ 的指数随机变量可表示为 $\exp(\lambda)$. □

使用记号 $\sim$ 表示分布, 这样 $X \sim \exp(\lambda)$ 表示 X 是具有平均值 $1/\lambda$ 的指数随机变量. 类似地, $X \sim \mathcal{N}(0, \Sigma)$ 表示期望值为零且协方差为 Σ 的 Gaussian 分布.

假定 $\mathbb{E}|X| < \infty$, 并且设 $\mathcal{G}$ 是 $\mathcal{F}$ 中的一个 σ 子代数, X 关于 $\mathcal{G}$ 的**条件期望**可以定义为函数 $\mathbb{E}[X|\mathcal{G}]: \Omega \mapsto E$, 并且满足

$$\int_G \mathbb{E}[X|\mathcal{G}]\,\mathrm{d}\mu = \int_G X\mathrm{d}\mu, \quad \forall G \in \mathcal{G},$$

也可以定义 $\mathbb{E}[f(X)|\mathcal{G}]$ 和条件概率 $\mathbb{P}[X \in F|\mathcal{G}] = \mathbb{E}[I_F(X)|\mathcal{G}]$, 其中 I_F 表示 F 的特征函数①.

3.3 随机过程

设 T 是一个有序集, 对于一个固定的 $t \in T, X_t$ 表示从 $(\Omega, \mathcal{F})$ 到 $(\mathrm{E}, \mathcal{G})$ 的随机变量, $X = \{X_t; t \in T\}$ 表示上述随机变量的集合, 称其为**随机过程**, 称测度空间 $\{\Omega, \mathcal{F}\}$ 为**样本空间**, $(\mathrm{E}, \mathcal{G})$ 表示**状态空间**. 在本书中, T 可表示为 $\mathbb{R}^+$ 或 $\mathbb{Z}^+$ ②, 则假定 T 中的所有元素都为非负. 状态空间 E 是定义在 $\mathbb{R}^d$ 上 Borel 集的 σ 代数, 有时也表示为 $\mathbb{T}^d$ 或者正整数子集 $\mathcal{I}$. 当 T 的上下文意思明了时, 也可以用 $\{X_t\}$ 代替 $\{X_t; t \in T\}$. 由于 X_t 为 $t \in T$ 和 $\omega \in \Omega$ 的函数, 因此, 有时为了方便起见, 用 $X(t), X(t, \omega)$ 或 $X_t(\omega)$ 代替 X_t.

随机过程的**有限维分布**是对于任意正整数 k 和 $t_i \in T\,(i \in \{1, \cdots, k\})$ 的 E^k 值随机变量 $(X(t_1), X(t_2), \cdots, X(t_k))$. 如果这样的随机变量组成的所有集合按照时间是同分布的, 则称这个过程为**稳定的**, 即对于任意整数 k, 时刻 $t_i \in T$ 和任意的 s, 满足 $s + t_i \in T(i \in \{1, \cdots, k\}), (X(t_1), X(t_2), \cdots, X(t_k))$ 的分布与 $(X(s+t_1), X(s+t_2), \cdots, X(s+t_k))$ 的分布是相同的.

定义 3.4 如果 $E = \mathbb{R}^d$, 并且所有的有限维分布都是 Gaussian(可能是退化的) 的, 则称之为 Gaussian 过程.

Gaussian 过程 $x(t)$ 是由它的平均数

$$m(t) := \mathbb{E}x(t)$$

① 这个函数在集合 F 上是 1, 在其他集合上是 0.

② 使用对流 $\mathbb{R}^+ = [0, \infty)$ 和 $\mathbb{Z}^+ = \{0, 1, 2, \cdots\}$.

和协方差函数

$$C(t,s)=\mathbb{E}((x(t)-m(t))\otimes(x(s)-m(s)))$$

来定义的, 此处对于固定的 $t,s\in T$, $m(t)\in\mathbb{R}^d$ 且 $C(t,s)\in\mathbb{R}^{d\times d}$.

由于 **Brownian 运动**是一种最重要的连续时间随机过程, 下面给出其定义, 首先是针对一维的, 然后再针对任意无限维 d.

定义 3.5　(i) 一维标准 Brownian 运动 $W(t):\mathbb{R}^+\to\mathbb{R}$ 是一个具有如下特性的实值随机过程:

a) $W(0)=0$;

b) $W(t)$ 是连续的;

c) $W(t)$ 的增量 $W(t)-W(s)$ 在非重叠的区间里是独立的, 而且对于每一个 $t>s\geqslant 0$, $W(t)-W(s)$ 都服从期望值为 0, 协方差为 $t-s$ 的 Gaussian 分布, 即随机变量 $W(t)-W(s)$ 的概率密度为

$$g(x;t,s)=(2\pi(t-s))^{-1/2}\exp\left(-\frac{x^2}{2(t-s)}\right). \tag{3.3.1}$$

(ii) d 维标准 Brownian 运动 $W(t):\mathbb{R}^+\to\mathbb{R}^d$ 是由 d 个独立的一维 Brownian 运动所构成的集合,

$$W(t)=(W_1(t),\cdots,W_d(t)),$$

其中 $W_i(t)\,(i=1,\cdots,d)$ 是独立的一维 Brownian 运动, 这样 Gaussian 随机向量 $W(t)-W(s)$ 的概率密度为

$$g(x;t,s)=(2\pi(t-s))^{-d/2}\exp\left(-\frac{\|x\|^2}{2(t-s)}\right).$$

有时称 Brownian 运动为 **Wiener 过程**, 它是一种 Gaussian 过程. 注意到对于 d 维 Brownian 运动, 设 I 表示 $d\times d$ 单位阵, 则有 (见 (3.2.1) 和 (3.2.2))

$$\mathbb{E}W(t)=0,\quad \forall t\geqslant 0$$

和

$$\mathbb{E}((W(t)-W(s))\otimes(W(t)-W(s)))=(t-s)I, \tag{3.3.2}$$

而且

$$\mathbb{E}(W(t)\otimes W(s))=\min(t,s)\,I. \tag{3.3.3}$$

下面将给出另外一个重要的连续时间过程 ——**Poisson 过程**. 注意到通过 (3.2.4) 定义的指数随机变量可得如下定义:

定义 3.6 具有 λ 的 Poisson 过程 $N(t)$ 是一个独立增量且满足

$$\mathbb{P}\left[(N(t)-N(s))=k\right]=\frac{\mathrm{e}^{-\lambda(t-s)}\left(\lambda(t-s)\right)^{k}}{k!},\quad t>s\geqslant 0, k\in\mathbb{N},$$

并且是整数值、连续时间的随机过程. Brownian 运动和 Poisson 过程都是齐次的, 或者是时间齐次的, 即在连续时间 s 和 t 之间的增量所满足的分布规律只依赖于 $t-s$.

设 $(\Omega,\mathcal{F},\mu)$ 是一个概率空间, (E,ρ) 是一个度量空间, 并且设 $T=[0,\infty)$. 设 $\{X_t\}$ 是从 $(\Omega,\mathcal{F},\mu)$ 到 (E,ρ) 具有连续样本路径的随机过程, 即对于所有的 $\omega\in\Omega$ 有 $X_t\in C_E:=C([0,\infty);E)$, 称这样的连续函数空间 C_E 为随机过程的样本空间. 在 C_E 上给定如下度量:

$$\rho_E\left(X^1,X^2\right):=\sum_{n=1}^{\infty}\frac{1}{2^n}\max_{0\leqslant t\leqslant n}\min\left(\rho\left(X_t^1,X_t^2\right),1\right).$$

通过这种度量所导出的拓扑可以定义 C_E 上的 Borel 集, 并且可以认为 $\{X_t\}$ 是定义在 $(\Omega,\mathcal{F},\mu)$ 上具有状态空间 $(C_E,\mathcal{B}(C_E))$ 的随机变量, 称定义在 $(C_E,\mathcal{B}(C_E))$ 上的概率测度 μX^{-1} 为 $\{X_t\}$ 的律, 则随机过程的律是它路径空间的概率测度. 在许多例子中, 随机过程极限定理的研究可归纳为在路径空间中对概率测度序列的分析.

例 3.7 连续函数空间 C_E 是 Brownian 运动 (Wiener 过程) 的路径空间, Brownian 运动的律是在 $C\left([0,\infty),\mathbb{R}^d\right)$ 上导出的测度, 称为 **Wiener 测度**. □

尽管 Brownian 运动是连续的, 然而许多随机过程并不具有连续的样本路径, 并且它们的路径空间比连续函数空间 C_E 要大, 用 D_E 表示右连续且具有左极限的过程 $X_t:[0,\infty)\mapsto E$ 所构成的空间①. 通常称 D_E 为 càdlàg 函数空间②, 称样本路径在 D_E 中的随机过程为 càdlàg 过程. 空间 D_E 在许多随机过程的极限定理中扮演着很有用的角色.

设 $(\Omega,\mathcal{F})$ 是一个测度空间, 并且 T 是一个有序集, 设 $X=X_t(\omega)$ 是建立在样本空间 $(\Omega,\mathcal{F})$ 到状态空间 $(E,\mathcal{G})$ 的一个随机过程, 它是两个变量 $t\in T$ 和 $\omega\in\Omega$ 的函数. 对一个固定的 $\omega\in\Omega$, 函数 $t\mapsto X_t(\omega)\,(t\in T)$ 是一个与 ω 有关的随机过程 X 中的样本路径. 设 $\mathcal{K}$ 是由 Ω 的子集所构成的集合, 用 $\sigma(\mathcal{K})$ 表示定义在 Ω 上的包含 $\mathcal{K}$ 的最小的 σ 代数, 有时也称之为由 $\mathcal{K}$ 所生成的 σ 代数. 设 $X_t:\Omega\mapsto E\,(t\in T)$, 映射集 $\{X_t,t\in T\}$ 是一个从样本空间 $(\Omega,\sigma(X_t,t\in T))$ 到状态空间 $(E,\mathcal{G})$ 的随机过程, 则其最小的 σ 代数 $\sigma(X_t,t\in T)$ 也称为由 $\{X_t,t\in T\}$ 所生成的 σ 代数.

---

① 即对于所有 $t\in[0,\infty)$, $\lim\limits_{s\to t^+}X_s=X_t$ 并且 $\lim\limits_{s\to t^-}X_s:=X_{t-}$ 存在.

② 该符号来自法语, 表示左连续、右极限的意思.

$(\Omega, \mathcal{F})$ 上的一种 $\mathcal{F}: \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}(s \leqslant t)$ 的子 σ 代数的非递减集 $\{\mathcal{F}_t, t \in T\}$, 并且设 $\mathcal{F}_\infty = \sigma\left(\bigcup_{t\in T} \mathcal{F}_t\right)$. 由 X_t 所生成的 σ 代数

$$\mathcal{F}_t^X := \sigma\left(X_s; s \leqslant t\right).$$

如果对于所有的 $t \in T$, X_t 是一个 $\mathcal{F}_t$ 测度的随机变量, 则称随机过程 $\{X_t; t \in T\}$ 是适合于 $\{\mathcal{F}_t\} := \{\mathcal{F}_t, t \in T\}$ 的.

定义 3.8　设 $\{X_t\}$ 是定义在概率空间 $(\Omega, \mathcal{F}, \mu)$ 上且在 E 中取值的随机过程, 并且设 $\{\mathcal{F}_t^X\}$ 是由 $\{X_t\}$ 所生成的, 如果对所有的 $t, s \in T, t \geqslant s$ 和 $\Gamma \in \mathcal{B}(E)$, $\{X_t\}$ 满足如下关系:

$$\mathbb{P}\left(X_t \in \Gamma | \mathcal{F}_s^X\right) = \mathbb{P}\left(X_t \in \Gamma | X_s\right), \tag{3.3.4}$$

则称之为 Markov 过程.

简单地说, 对于任意的 $t \geqslant s$, 一旦 X_s 已知, 则就可完全决定 X_t 的统计. 因此, $X_t (t < s)$ 的信息是多余的, 即 Markov 过程不具有记忆性.

为了介绍 Markov 过程的性质, 先假设存在函数 $P: T \times T \times E \times \mathcal{B}(E) \to \mathbb{R}^+$ 满足 Chapman-Kolmogorov 方程:

$$\int_E P\left(s, t, x, \mathrm{d}y\right) P\left(t, v, y, \Gamma\right) = P\left(s, v, x, \Gamma\right) \tag{3.3.5}$$

对于所有的 $x \in E, \Gamma \in \mathcal{B}(E)$ 和 $s, t, v \in T\,(s \leqslant t \leqslant v)$. 如果 P 满足 Chapman-Kolmogorov 方程以及对于所有的 $s, t \in T\,(t \geqslant s)$, $\Gamma \in \mathcal{B}(E)$, 满足如下关系:

$$\mathbb{P}\left(X_t \in \Gamma | \mathcal{F}_s^X\right) = P\left(s, t, X_s, \Gamma\right) \quad \text{a.s.},$$

则称随机过程 $\{X_t\}$ 是带有转移函数 P 的 Markov 过程①.

如果 $P(s, t, x, \Gamma) = P(0, t - s, x, \Gamma)$, 则称 Markov 过程是齐次的, 简化记号, 设 $P(0, t, \cdot, \cdot) = P(t, \cdot, \cdot)$, 则 Chapman-Kolmogorov 方程变为

$$\int_E P\left(s, x, \mathrm{d}y\right) P\left(t, y, \Gamma\right) = P\left(t + s, x, \Gamma\right). \tag{3.3.6}$$

例 3.9　一维 Brownian 运动是齐次的 Markov 过程, 如果转移函数是例 3.2 中定义的 Gaussian, 则有

$$P\left(t, x, \mathrm{d}y\right) = \gamma_{t,x}(y) \mathrm{d}y.$$

□

例 3.10　Poisson 过程是齐次的 Markov 过程.　□

① 见定义 3.6 中的讨论.

设 (E,ρ) 是一个度量空间, 并且 $\{X_t\}$ 是一个 E 值齐次 Markov 过程. 设 $T=\mathbb{R}^+$, 对于任意 $f\in C_b(E)$, 通过如下方式:

$$\mathcal{T}(t)f(x)=\int f(y)P(t,x,\mathrm{d}y)=\mathbb{E}\left(f\left(X_t\right)|X_0=x\right)$$

定义一个参数算子集 $\mathcal{T}(t)$, 并且设 $\mathcal{T}(t):C_b(E)\to C_b(E)$, 则 $\mathcal{T}(t)$ 形成在 $C_b(E)$ 上的算子半群. 用 $\mathcal{D}(\mathcal{L})$ 表示 $f\in C_b(E)$ 且在 $C_b(E)$ 中的强收敛, 即

$$\mathcal{L}f=\lim_{t\to 0}\frac{\mathcal{T}(t)f-f}{t}$$

的函数所组成的集合. 称算子 $\mathcal{L}:\mathcal{D}(\mathcal{L})\to C_b(E)$ 为算子半群 $\mathcal{T}(t)$ 的**无穷小生成子**. 相应地, 有如下定义:

定义 3.11 称先前定义的算子 $\mathcal{L}:C_b(E)\to C_b(E)$ 为 Markov 过程 $\{X_t\}$ 的生成子.

在概率背景下, 出现空间 $C_b(E)$ 是自然的, 然而, 在实际中也会经常出现一些其他的 Banach 空间. 例如, 当 E 中存在度量 μ 时, 有时就会产生空间 $L^p(E;\mu)$(见第 4, 5 章). 然而, 在本章中研究易于表示的 $C_b(E)$ 空间.

通常将生成子看成是定义齐次 Markov 过程的起点. 相反地, 令 $\mathcal{T}(t)$ 是定义在 $\mathcal{D}(\mathcal{T}(t))\subset C_b(E)$ 上具有闭合的压缩算子半群①, 那么在相应的假设下, 存在带有压缩算子半群的 E 值齐次 Markov 过程 $\{X_t\}$, 其中 $\mathcal{T}(t)$ 满足

$$\mathbb{E}\left(f\left(X(t)\right)|\,\mathcal{F}_s^X\right)=\mathcal{T}\left(t-s\right)f\left(X(s)\right),$$

其中 $t,s\in T$ 且 $t\geqslant s, f\in\mathcal{D}\left(\mathcal{T}(t)\right)$.

例 3.12 再次考虑例 3.9 中的一维 Brownian 运动. 与标准 Brownian 运动相关的半群是热半群 $\mathcal{T}(t)=\exp\left(\frac{t}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)$, 这个 Markov 过程的生成子是 $\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$. □

遍历性在研究随机过程极限定理的过程中是一个非常重要的概念. 在 Markov 过程中, 遍历性提供了有关 Markov 半群的长时间行为信息.

定义 3.13 如果等式

$$\mathcal{T}(t)g=g,\quad g\in C_b(E),\ \forall t\geqslant 0$$

有唯一常数解, 则称 Markov 过程具有**遍历性**.

大体上说遍历性是对应如下情形的, 即半群 $\mathcal{T}(t)$ 使得 $\mathcal{T}(t)-I$ 在其零空间上有唯一常数, 相当于生成子 $\mathcal{L}$ 在其零空间上有唯一常数. 然而, 当以这样的方式制

① 设 X 是一个 Banach 空间, 并且令 $T:X\to X$ 是一个有界算子. 如果 $\|Tf\|_X\leqslant\|f\|_X$ $(\forall f\in X)$, 则称 T 是一个压缩.

定遍历性, 尤其是针对确定性动力系统时, 如何选择函数空间就是一个敏感的问题 (见第 4 章).

在附加假设下, 当 $T=\mathbb{R}^+$ 时, 遍历性 Markov 过程具有如下性质:

$$\lim_{t\to+\infty}\frac{1}{t}\int_0^t g(X_s)\mathrm{d}s\to\mathbb{E}g(x),$$

则称 Markov 过程在 $\mathbb{E}$ 上有不变的测度 μ, 其中 $\mathbb{E}$ 定义为关于 μ 的期望. 对于初始点 X_0 的测度以及基本 Markov 过程的概率测度可能会出现不同的收敛形式, 而且如果 X_0 是服从 μ 的分布, 那么 $X_t\,(t>0)$ 均是服从 μ 的分布. 这种由服从 X_0 分布得到的随机过程是稳态的.

正如前面所提到的, 有时将随机过程 X_t 看成是含有变量 $t\in T$ 以及 $\omega\in\Omega$: $X(t,\omega)$ 的函数是有用的. 在这种背景下, 正如在前面章节中提到过的 (见定义 2.22 及其随后的讨论) 一样, 关注 Banach 值空间是有趣的.

例 3.14　考虑在时值连续函数 $E=C([0,T],\mathbb{R})$ 空间中的随机过程 X, 并令 $p\in[1,\infty)$, 定义 $L^p(\Omega,E)$ 为 Banach 空间, 该空间赋予的范数 $\|\cdot\|_{L^p(\Omega,E)}$ 为

$$\|X\|^p_{L^p(\Omega,E)}=\mathbb{E}\left(\left(\sup_{t\in[0,T]}|X(t)|\right)^p\right).$$

注意: 此定义等价于

$$\|X\|^p_{L^p(\Omega,E)}=\mathbb{E}\left(\sup_{t\in[0,T]}|X(t)|^p\right),$$

并且通常范数以这样的形式表示. □

3.4　鞅和随机积分

3.4.1　鞅

定义 3.15　令 $\{\mathcal{F}_t\}$ 是概率空间 $(\Omega,\mathcal{F},\mu)$ 上的 σ 代数, $\{M_t\}$ 为 $\mathbb{R}^d$ 值随机过程, 该随机过程服从 $\{\mathcal{F}_t\}$ 并满足 $\mathbb{E}|M_t|<\infty(t\in T)$. 如果

$$\mathbb{E}[M_t|\,\mathcal{F}_s]=M_s,\quad t,s\in T,t>s,$$

则称 $\{M_t\}$ 为 $\mathcal{F}_t$ 鞅.

那么对于鞅

$$\mathbb{E}[M_t-M_t|\,\mathcal{F}_s]=0,\quad t,s\in T,t>s$$

是以 $\mathcal{F}_s$ 为条件的, 其增量平均值为 0.

如果 $\mathbb{E}|M_t|^2 < \infty\,(t \in T)$, 那么鞅 M_t 是平方可积的.

例 3.16 通过定义 3.5, 标准 d 维 Brownian 运动是一个由本身生成的 σ 代数的鞅. □

在随机过程理论及其应用中, 鞅的二次变差是很重要的, 为了解释它需要如下定义:

定义 3.17 在 $t \in [0,T]$ 上定义的 $\mathbb{R}^{d\times d}$ 值过程 $Q(t)$, 其中 $Q(0)=0$. 当 $Q(t)\,(t\in[0,\infty))$ 非负并且 $Q(t)-Q(s)\,(0\leqslant s\leqslant t\leqslant T)$ 非负时, $Q(t)$ 是增长的.

定义 3.18 如果

$$M(t)\otimes M(t)-Q(t)$$

是 $\mathcal{F}_t$ 鞅, 则 $\mathbb{R}^{d\times d}$ 值随机增过程 $Q(t)$ 是 $\mathbb{R}^d$ 值 $\mathcal{F}_t$ 鞅 $M(t)$ 的二次变分, 记 $Q(t)=\langle M\rangle_t$.

定理 3.19 令 $M(t)$ 为连续的、$\mathbb{R}^d$ 值、平方可积的 $\mathcal{F}_t$ 鞅, 则其一定有一个平方变差过程.

例 3.20 (i) 一维标准 Brownian 运动的二次变分为 t. 注意到 $M(t):=W(t)^2-t$ 是一个鞅, 更一般地说, 令 $W(t)$ 是 $\mathbb{R}^d$ 上的标准 Brownian 运动, 那么 $(W(t)\otimes W(t)-It)$ 是一个鞅, 因此, $\langle W\rangle_t=It$.

(ii) 令 X_t 是生成子为 $\mathcal{L}$ 的 $\mathbb{R}$ 值 Markov 过程, 并且 $\phi\in\mathcal{D}(\mathcal{L})$, 其中 $\mathcal{D}(\mathcal{L})$ 为 $\mathcal{L}$ 的定义域, 则

$$M_t=\phi(X_t)-\phi(X_0)-\int_0^t(\mathcal{L}\phi)(X_s)\,\mathrm{d}s \tag{3.4.1}$$

是一个鞅. 设 $\phi^2\in\mathcal{D}(\mathcal{L})$, 则由如下公式得到 M_t 的二次变差:

$$\langle M_t\rangle=\int_0^t((\mathcal{L}\phi^2)(X_s)-2\phi(X_s)(\mathcal{L}\phi)(X_s))\mathrm{d}s. \tag{3.4.2}$$

(iii) 采用前面两个例子中的记号, 令 $\phi(x)=x, X_t=W(t)$, 并且 $d=1$, 则 $\mathcal{L}\phi(x)=0, \mathcal{L}\phi^2(x)=1, M_t=W(t)$ 且 $\langle W\rangle_t=t$. □

如果 $\lambda>0, p\geqslant 1, t\geqslant 0$, 则每个连续的鞅满足如下 **Doob 鞅不等式**:

$$\mathbb{P}\left[\sup_{0\leqslant s\leqslant t}|M_s|\geqslant\lambda\right]\leqslant\frac{1}{\lambda^p}\mathbb{E}|M_t|^p. \tag{3.4.3}$$

此外, 当 $p>1, t\geqslant 0$ 时,

$$\mathbb{E}\left(\sup_{0\leqslant s\leqslant t}|M_s|^p\right)\leqslant\left(\frac{p}{p-1}\right)^p\mathbb{E}|M_t|^p. \tag{3.4.4}$$

3.4.2 Itô随机积分

对随机微分方程进行严密分析时, 有必要定义如下随机积分形式:

$$I(t) = \int_o^t f(s)\mathrm{d}W(s), \tag{3.4.5}$$

其中 $W(t)$ 是 d 维 Brownian 运动, 并且 $f(s) \in \mathbb{R}^{m\times d}$. 因为 $W(t)$ 没有有界变差路径, 使得问题不是很明晰. 先以 $m = d = 1$ 举例, 为了定义随机积分, 假设 $f(t)$ 是随机过程, 该随机过程适于由 $W(t)$ 得到的 $\mathcal{F}(t)$, 那么

$$\mathbb{E}\left(\int_0^T f(s)^2\mathrm{d}s\right) < \infty.$$

下面用 Itô方法来说明随机积分. Itô随机积分 $I(t)$ 可用式 (3.4.5) 定义的 Riemann 近似求和的 L^2 极限来表示,

$$I(t) := \lim_{K\to\infty}\sum_{k=1}^{K-1} f\left(t_{k-1}\right)\left(W\left(t_k\right) - W\left(t_{k-1}\right)\right), \tag{3.4.6}$$

其中 $t_k = k\Delta t$, 并且 $K\Delta t = t$. 注意在 (3.4.6) 中, 函数 f 值是由每个区间 $[t_{n-1}, t_n]$ 的左端点求得的. Itô随机积分 $I(t)$ 关于 t 是连续的, 当 $W(s)$ 是标准 d 维 Brownian 运动, 并且对每个 $s, f(s) \in \mathbb{R}^{m\times d}$ 时, 导出的过程 $I(t) \in \mathbb{R}^m$, 此时这些想法可以很容易地归纳为这种例子.

积分结果满足 Itô等距映射

$$\mathbb{E}\left|I(t)\right|^2 = \int_0^t \mathbb{E}\left|f(s)\right|_F^2\mathrm{d}s, \tag{3.4.7}$$

其中 $|\cdot|_F$ 表示由式 (2.2.1) 定义的 Frobenius 范数, 而且这个积分是一个鞅,

$$\mathbb{E}I(t) = 0,$$

并且

$$\mathbb{E}\left[I(t)|\,\mathcal{F}_s\right] = I(s), \quad \forall t \geqslant s,$$

其中定义 $\mathcal{F}_s$ 为由 $W(s)$ 生成的.

例 3.21 考虑 Itô随机积分

$$I(t) = \int_0^t f(s)\mathrm{d}W(s),$$

其中 f 和 W 是标量值. 这是一个具有二次变差的鞅,

$$\langle I\rangle_t = \int_0^t (f(s))^2\mathrm{d}s.$$

通常对于任意有限维的 f 和 W, 积分 $I(t)$ 是具有二次变差的鞅,

$$\langle I\rangle_t = \int_0^t (f(s)\otimes f(s))\mathrm{d}s. \qquad \square$$

注意到

$$\mathbb{E}\left|\langle I\rangle_t\right| \leqslant \int_0^t \mathbb{E}\left|f(s)\right|^2\mathrm{d}s \leqslant C\int_0^t \mathbb{E}\left|f(s)\right|_F^2\mathrm{d}s = C\mathbb{E}\left|I(t)\right|^2,$$

相似于二阶变差过程矩的有界性, 对任意的 $p\geqslant 1$, 假定其他高阶矩的有界性 $|f|_F$,

$$\mathbb{E}(|\langle I\rangle_t|^{p/2}) \leqslant C. \tag{3.4.8}$$

鞅的一个最重要的特性就是它满足不同路径的不等式, 当证明极限定理时, 这些不等式尤为重要, 随后将介绍一种最重要的鞅不等式.

定理 3.22 (Burkholder-Davis-Gundy 不等式) 考虑 Itô随机积分 (3.4.5), 二次变差的鞅 $\langle I\rangle_t$, 对于所有 $p>0$, 存在一个正常数 $C_p^{\pm}$, 使得

$$C_p^-\mathbb{E}\left(\left|\langle I\rangle_t\right|^{p/2}\right) \leqslant \mathbb{E}\left(\sup_{0\leqslant s\leqslant t}|I(s)|^p\right) \leqslant C_p^+\mathbb{E}\left(\left|\langle I\rangle_t\right|^{p/2}\right).$$

由 Itô公式 (见第 6 章) 以及 Doob 鞅不等式可对定理 3.22 进行证明.

3.4.3 Stratonovich 随机积分

除了 Itô随机积分, 下面的 Stratonovich 随机积分有时也很有用. 该积分定义为对式 (3.4.5) 进行 Riemann 差分求和时的近似 L^2 极限, 即

$$I_{\text{strat}}(t) := \lim_{K\to\infty}\sum_{k=1}^{K-1}\frac{1}{2}\left(f\left(t_{k-1}\right)+f\left(t_k\right)\right)\left(W\left(t_k\right)-W\left(t_{k-1}\right)\right), \tag{3.4.9}$$

其中 $t_k = k\Delta t$ 以及 $K\Delta t = t$. 注意在 (3.4.9) 中, 函数 $f(t)$ 的值是由每个区间 $[t_{n-1},t_n]$ 的两端点取得的. 由类似方式, 也可以定义多维 Stratonovich 积分, 积分结果记为

$$I_{\text{strat}}(t) = \int_0^t f(s)\circ\mathrm{d}W(s).$$

由 (3.4.9) 中的极限得到一种不同于 Itô积分的积分形式. 这样对于有界变差函数, 积分情形比由标准 Riemann 积分理论得出的积分更为复杂. 那样的话, 通过极限法被积函数取值区间 $[t_{k-1},t_k]$ 中的点就不会影响积分的定义. 至于对不具有界变差性质的 Brownian 运动进行逆向积分时, 极限就变得不同了. 然而, 当通过随机微分方程弄清楚 f 与 W 之间的关联时, 这里存在一个转换公式, 见第 6 章. 一般来说, 通过积分类来定义 Stratonovich 积分要比定义 Itô积分困难.

3.5 概率测度的弱收敛

在概率论中常用到的一类收敛是**概率测度的弱收敛**.

定义 3.23　令 (E,ρ) 是具有 Borel σ 代数 $\mathcal{B}(E)$ 的度量空间, 令 $\{\mu_n\}_{n=1}^{\infty}$ 是 $(E,\mathcal{B}(E))$ 上的概率测度序列, 并且令 μ 是此空间上的另一个测度. 如果对于 $f\in C_b(E)$ 有

$$\lim_{n\to\infty}\int_E f(s)\mathrm{d}\mu_n(s)=\int_E f(s)\mathrm{d}\mu(s),$$

则称 $\{\mu_n\}_{n=1}^{\infty}$ 弱收敛于 μ, 并记 $\mu_n\Rightarrow\mu$.

定义 3.24　令 $\{\Omega_n,\mathcal{F}_n,\mu_n\}_{n=1}^{\infty}$ 是概率空间上的序列且 (E,ρ) 是一个度量空间. 令 $X_n:\Omega_n\mapsto E\,(n=1,2,\cdots)$ 是一个随机变化的序列. 假设 $(\Omega,\mathcal{F},\mu)$ 是另一个概率空间且 $X:\Omega\mapsto E$ 是一个随机变化的序列. 如果在 $(E,\mathcal{B}(E))$ 上的测度序列 $\{\mu X_n^{-1}\}_{n=1}^{\infty}$ 弱收敛于 $(E,\mathcal{B}(E))$ 上的 μX^{-1}, 则称 $\{X_n\}_{n=1}^{\infty}$ 依分布或弱收敛于 X, 并记 $X_n\Rightarrow X$.

换言之, 对于 $f\in C_b(E)$, 当且仅当

$$\lim_{n\to\infty}\mathbb{E}_nf(X_n)=\mathbb{E}f(X)$$

时, $X_n\Rightarrow X$, 其中 $\mathbb{E}_n,\mathbb{E}$ 分别定义为 μX_n^{-1} 以及 μX^{-1} 下的期望. 下面的例子说明了在概率论中弱收敛的重要性.

例 3.25 (中心极限定理)　令 $\{\xi_n\}_{n=1}^{\infty}$ 是 i.i.d 的随机变化的序列, 其均值为零, 并且方差为 1. 定义

$$S_n:=\sum_{k=1}^{n}\xi_k, \tag{3.5.1}$$

则序列

$$X_n:=\frac{1}{\sqrt{n}}S_n$$

依分布收敛于标准正态随机变量. □

在研究随机变量以及随机过程的极限理论时, 也会用到其他不同的收敛类型. 在概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上, 令 $\{X_n\}_{n=1}^{\infty}$ 为 $(E,\mathcal{B}(E))$ 的随机变化序列, 而 X 为另一个随机变量. 假定 E 是 Banach 空间, 并且其范数用 $\|\cdot\|$ 表示. 如果对于 $\varepsilon>0$ 有

$$\lim_{n\to\infty}\mathbb{P}(\|X_n-X\|>\varepsilon)\to 0,$$

则称序列**依概率收敛**于随机变量 X.

如果

$$\mathbb{P}\left(\lim_{n\to\infty}X_n=X\right)=1,$$

则称 $\{X_n\}_{n=1}^{\infty}$ **几乎是收敛的**(或者说是概率为 1).

最后, 假如

$$\lim_{n\to\infty} \mathbb{E}\,\|X_n - X\|^p = 0,$$

则称序列 $\{X_n\}_{n=1}^{\infty}$ **依p阶均值** (或者是在 L^p 上)**收敛**.

注意这仅仅是对于 Ω 中的 μ 积分而言, 得到了 Banach 空间 $L^p(\Omega; E)$ 上的强收敛.

例 3.26 (强大数定理) 令 $\{\xi_n\}_{n=1}^{\infty}$ 是均值为 m 的 i.i.d 序列, 定义

$$X_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k,$$

则序列 X_n 一定收敛于 m. 进而假定 $\xi_n \in L^2$, 则 X_n 在 L^2 上也收敛于 1.

注 3.27 在具有不同定义的随机变量收敛之间存在一些有用的关系. 例如, 强收敛中包含弱收敛, 概率收敛隐含了弱收敛. □

接下来继续讨论弱收敛. 特别地, 在连续映射条件下, 弱收敛是其固有的一种性质.

定理 3.28 令 $\{\Omega_n, \mathcal{F}_n, \mu_n\}_{n=1}^{\infty}$ 为概率空间上的一个序列, $(E_i, \rho_i)(i = 1, 2)$ 是度量空间, $X_n : \Omega_n \mapsto E_1\,(n = 1, \cdots, \infty)$ 是随机变量序列. 假定 $(\Omega, \mathcal{F}, \mu)$ 是另一个概率空间, 并且令 $X : \Omega \mapsto E_1$ 为另一个随机变量. 如果 $f : E_1 \to E_2$ 是连续的, $\{X_n\}_{n=1}^{\infty}$ 依分布收敛于 X, 则 $\{f(X_n)\}_{n=1}^{\infty}$ 依分布收敛于 $f(X)$.

例 3.29 令 $E_1 = C([0,1], \mathbb{R})$ 且 $E_2 = \mathbb{R}$, 由

$$f(x) := \sup_{t\in[0,1]} x(t)$$

定义的函数 $f : E_1 \to E_2$ 是连续的. 因此, 根据定理 3.28, 在 $C([0,1], \mathbb{R})$ 上的 $X_n \Rightarrow X$ 暗含了在 $\mathbb{R}$ 上 $\sup_{t\in[0,1]} X_n(t) \Rightarrow \sup_{t\in[0,1]} X(t)$. □

通常, 当 $X_n \Rightarrow X$ 以及 $Y_n \Rightarrow Y$ 时, (X_n, Y_n) 间的相互关系使得对偶 (X_n, Y_n) 不会弱收敛于 (X, Y). 然而, 如果极限 Y 不是随机的 (有关随机变量 Y 的概率测度是点式群体), 则对偶就不会收敛.

定理 3.30 如果 $X_n \Rightarrow X$ 且 $Y_n \Rightarrow Y$, 其中 Y 是非随机的, 则 $(X_n, Y_n) \Rightarrow (X, Y)$.

设 $(\Omega, \mathcal{F}, \mu)$ 为概率空间, $T = \mathbb{R}^+$ 以及 (E, ρ) 是度量空间. 令 $\{X_t^n\}_{n=1}^{\infty}$ 是随机过程的一个族, 而 X_t 为另一个随机过程, 它们在 $C_E = C([0, \infty), E)$ 上都有样本路径. 回想如果在 $(C_E, \mathcal{B}(C_E))$ 上, 概率测度序列 $\mathbb{P}(X^n)^{-1}$ 弱收敛于概率测度 $\mathbb{P}X^{-1}$ 时, 那么序列 $\{X^n\}_{n=1}^{\infty}$ 弱收敛于 X 且记为 $X^n \Rightarrow X$. 有时称 $\{X_t^n\}_{n=1}^{\infty}$ 在 C_E 上弱收敛, 或者说, $\{X_t^n\}_{n=1}^{\infty}$ 在路径空间上是弱收敛的.

例 3.31 (函数中心极限定理 —— 不变性原理) 考虑例 3.25 中的情形和由 (3.5.1) 给出的 S_n. 用 $[t]$ 表示实数 t 的整数部分, 并且定义如下连续时间过程:

$$X_t^n := \frac{1}{\sqrt{n}} S_{[nt]} + \frac{1}{\sqrt{n}} (nt - [nt])\, \xi_{[nt]+1}. \tag{3.5.2}$$

此过程有连续路径, 而且在 $C_{\mathbb{R}}$ 上, $\{X_t^n\}_{n=1}^{\infty}$ 弱收敛于标准一维 Brownian 运动. □

在例 3.25 和例 3.31 中, 基于独立随机变量求和, 分别讨论了典型的中心极限定理以及函数中心极限定理 (不变性原理). 如果随机变量之间的依赖性不是很强, 则对于相关联的随机变量也会得到类似的结果. 在很多例子中, 对独立性进行很好地概括的鞅性质能充分使极限定理成立, 下面的定理展示了这一性质, 并将先前的例子作为一个特例来说明它.

定理 3.32 设 $(\Omega, \mathcal{F}, \mu)$ 是一个概率空间且 $\{\mathcal{F}_j, j \geqslant 1\}$ 是一个过滤. 令 $\{\xi_j, j \geqslant 1\}$ 为适应的遍历性过程, 满足

$$\mathbb{E}[\xi_1]^2 = \sigma^2$$

和

$$\mathbb{E}\left[\xi_{k+1} \,|\, \mathcal{F}_k\right] = 0. \tag{3.5.3}$$

定义

$$S_n := \sum_{k=1}^{n} \xi_k, \tag{3.5.4}$$

则

$$X^n := \frac{1}{\sqrt{n}} S_n$$

依分布收敛于一个 Gaussian 变量, 其均值为 0, 方差为 σ^2, 而且过程

$$X_t^n := \frac{1}{\sqrt{n}} S_{[nt]} + \frac{1}{\sqrt{n}} (nt - [nt])\, \xi_{[nt]+1}$$

在 $C_{\mathbb{R}}$ 上弱收敛于 $\sigma W(t)$, 其中 $W(t)$ 为标准 Brownian 运动. 最后, 过程

$$X_t^n := \frac{1}{\sqrt{n}} S_{[nt]}$$

在 $D_{\mathbb{R}}$ 上弱收敛于 $\sigma W(t)$.

注意条件 (3.5.3) 暗含 (3.5.4) 中定义的 S_n 是一个 $\mathcal{F}_j$ 鞅. 对于任意维 $d \geqslant 1$, 前面的定理一样适用, 即 $\{\xi_j, j \geqslant 1\}$ 是一个不变的、具有遍历性随机变量的向量值序列. 此时, 极限过程是一个标准的 d 维 Brownian 运动右乘协方差阵 ξ_1 的平方根之积.

对于连续时间鞅可以证明一个结论, 该结论类似于定理 3.32, 接下来说明它.

定理 3.33 (鞅中心极限定理) 令 $\{M(t):\mathbb{R}^+\mapsto\mathbb{R}^d\}$ 是关于 $\{\mathcal{F}_j: t\geqslant 0\}$ 的概率空间 $(\Omega,\mathcal{F},\mu)$ 上的连续平方积分鞅, 设 $\langle M\rangle_t$ 为其二次变差过程. 假设

(i) $M(0)=0$;

(ii) 过程 $M(t)$ 拥有连续样本路径且其增量不变;

(iii) 在 $L^1(\mu)$ 上 $M(t)$ 的二次变差收敛于某个对称正定矩阵 Σ,

$$\lim_{t\to\infty}\mathbb{E}\left(\left|\frac{\langle M\rangle_t}{t}-\Sigma\right|\right)=0. \tag{3.5.5}$$

于是过程 $1/\sqrt{t}M_t$ 依分布收敛于一个 $\mathcal{N}(0,\Sigma)$ 的随机变量, 而且变尺度鞅

$$M^{\varepsilon}(t):=\varepsilon M\left(\frac{t}{\varepsilon^2}\right)$$

在 $C_{\mathbb{R}^d}$ 上弱收敛于 $\sqrt{\Sigma}W(t)$, 其中 $W(t)$ 为标准的 d 维 Brownian 运动, $\sqrt{\Sigma}$ 为矩阵 Σ 的平方根.

大体上说, 鞅中心极限定理将二次变差的强大数定理与过程本身的波动性联系起来. 综合随机积分是鞅以及进行遍历性假设两方面的事实, 得到下面的推论.

推论 3.34 令 $W(t)$ 为标准的 d 维 Brownian 运动, 并且设 $z(t):\mathbb{R}^+\mapsto\mathbb{R}^d$ 是一个连续的、静止的、遍历性 Markov 过程, 该过程有不变测度 $\mu(dz)$ 并适用于由 $W(t)$ 生成的. 如果 $F:\mathbb{R}^d\mapsto\mathbb{R}^{m\times d}$ 是一个光滑有界函数. 定义

$$I(t)=\int_0^t F(z(s))\mathrm{d}W(s),$$

通过遍历性, 注意到在 $L^1(\mu)$ 上有

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t(F(z(s))\otimes F(z(s)))\mathrm{d}s=\int_{\mathcal{Z}}(F(z)\otimes F(z))\mu(\mathrm{d}z), \tag{3.5.6}$$

则对所有 $T>0$, 变尺度随机积分

$$I^{\varepsilon}(t):=\varepsilon I\left(\frac{t}{\varepsilon^2}\right)$$

依分布收敛于 $\sqrt{\Sigma}W(t)$, 其中 $W(t)$ 是一个 m 维 Brownian 运动, 并且

$$\Sigma=\int_{\mathbb{R}^d}(F(z)\otimes F(z))\,\mu(\mathrm{d}z). \tag{3.5.7}$$

证明 为了证明上述结论, 必须确保鞅 $I(t)$ 满足定理 3.33 中的假设, 显然有 $I(0)=0$. 而且 $z(t)$ 过程的平稳性暗含 $I(t)$ 有独立的增量. 通过 (3.5.6) 及 (3.5.7) 可得

$$\lim_{t\to\infty}\frac{1}{t}\langle I\rangle_t=\lim_{t\to\infty}\frac{1}{t}\int_0^t(F(z(s))\otimes F(z(s)))\mathrm{d}s=\int_{\mathcal{Z}}(F(z)\otimes F(z))\mu(\mathrm{d}z)=\Sigma.$$

因此, 根据定理 3.33, 结论成立. □

3.6 讨论和参考

本章中大部分的内容是非常标准的, 并且可在多本教科书中找到. 有关概率论的教科书见文献 [36, 50, 102, 103, 194, 195, 302]; 随机过程理论方面的教科书可见文献 [156, 270, 304].

有关概率测度及其性质的一般参考见文献 [252]; [37] 是一本关于概率测度弱收敛方面的优秀参考书籍; 对于 Markov 过程序列的各种收敛结果见文献 [94]; 有关路径空间中概率测度的随机过程和弱收敛方面的背景资料见文献 [120]; 对不同收敛模式之间关系的讨论见文献 [130]; 关于随机过程的极限理论见文献 [152]; 概率论和随机过程理论的应用在文献 [64] 中进行了很好的介绍.

Hille-Yosida 理论给出了一个算子成为连续半群生成子的充要条件[98,337]; 在 19 世纪 50 年代, Feller 给出了基于半群理论 Markov 过程的系统研究[79,196,337]; 使用半群算子极限定理证明不同 Markov 过程的极限理论 (如 **Trotter-Kato 定理**[157]), 这类问题已经在文献 [181] 中证明, 见文献 [94].

具有 L^2 空间 (由不变测度加权得到) 上对称生成子的遍历性 Markov 过程被称为是可逆的, 有关这方面讨论见文献 [303, ch.5]; 可逆 Markov 过程的对称生成子在证明极限定理中是很有用的, 见文献 [74, 127, 173].

定理 3.22 中的 Burkholder-Davis-Gundy 不等式不仅是针对于 Itô随机积分有相应的公式, 而且对任意鞅都有相应的公式[156,ch.3,Theorem3.28]; 二次变差矩 (如 (3.4.8)) 以及随机积分矩的边界见文献 [156, 210].

对例 3.20 结论的证明见文献 [270]. 出现在式 (3.4.2) 右边的算子有时也称为 Operateur carré du champ; 在文献 [94, ch.7] 和 [185, 186] 中可以找到鞅中心极限定理的证明过程; 大量关于鞅极限定理的详细介绍可见文献 [137].

在一维情形时, 连续局部鞅可以用时变 Brownian 运动表示, 有时称其为 **Dambis-Dubins-Schwarz 定理**[156,ch.3,Theorem4.6]. 令 $M=\{M_t,\mathcal{F}_t;0\leqslant t<\infty\}$ 是 $\mathbb{R}$ 值鞅, 满足

$$\lim_{t\to\infty}\langle M\rangle_t=\infty, \mathbb{P}\text{a.s.}$$

对于 $0\leqslant s<\infty$, 定义停止时间

$$T(s)=\inf\left\{t\geqslant 0;\langle M\rangle_t>s\right\},$$

则由下式:

$$B_s=M_{T(s)},\ \mathcal{G}_s=\mathcal{F}_{T(s)},\quad 0\leqslant s<\infty$$

给出的时变过程 B_s 以及 $\mathcal{G}_s$ 是一个标准的一维 Brownian 运动. 特别在 $\mathbb{P}$a.s.,

$$M_t=B_{\langle M\rangle_t}.$$

利用这个性质可以描述一系列鞅的极限, 见文献 [40, 134]. 注意到通常 B_t 以及 M_t 是高度相关的, 由于时间的改变使得它们不适用于同一种.

鞅中心极限定理引出对 Markov 过程可加函数的一般中心极限定理：令 $y(t)$ 是一个在 $\mathcal{Y}$ 上具有遍历性的 Markov 过程, 其生成子是 $\mathcal{L}$, 并且不变测度为 $\mu(\mathrm{d}y)$. 考虑下面的积分 ($y(t)$ 是一个可加函数):

$$x(t) = x_0 + \int_0^t f(y(s))\mathrm{d}s.$$

假设

$$\int_{\mathcal{Y}} f(y)\mu(\mathrm{d}y) = 0,$$

并且在某种恰当的意义下, Poisson 方程

$$-\mathcal{L}\phi = f \tag{3.6.1}$$

是适定的, 则变尺度过程

$$x^{\varepsilon}(t) = \varepsilon x\left(\frac{t}{\varepsilon^2}\right)$$

弱收敛于 $\sqrt{\Sigma}W(t), W(t)$ 是一个标准 Brownian 运动. 极限 Brownian 运动的方差由生成子 $\mathcal{L}$ 的 Dirichlet 形式给出 (见第 6 章), 可由 (3.6.1) 的解计算, 即

$$\Sigma = \int_{\mathcal{Y}} (-\mathcal{L}\phi) \otimes \phi \, \mu\,(\mathrm{d}y)\,.$$

在一维条件时, 注意 (3.4.1) 暗含了变尺度过程 $x^{\varepsilon}(t)$ 满足

$$x^{\varepsilon}(t) = \varepsilon(x_0 - \phi\left(x\left(\frac{t}{\varepsilon^2}\right)\right) + \phi(x(0))) + \varepsilon M\left(\frac{t}{\varepsilon^2}\right),$$

其中 $M(t)$ 是一个鞅. 假设 ϕ 满足恰当的估计, 则等式右边首项趋于 0. 依据鞅中心极限定理, 等式右边第二项收敛于方差为 Σ 的 Brownian 运动 (用不变测度解决这个问题时, 要求使用分部积分). 对于可逆 Markov 过程中这个定理的证明见文献 [74, 173]. 证明基于鞅极限定理的 Markov 过程极限定理方法在文献 [243] 中有进一步的论述. 对于这个结论的不同推论已经得到证明, 尤其是在对生成子作适当假设后, Markov 变为不可逆过程时的情形, 见文献 [186] 及其参考文献.

3.7 练　　习

1. 证明例 3.2 中关于 Gaussian 随机变量的均值和协方差的结论.
2. 证明 (3.3.2) 和 (3.3.3) 关于 Wiener 过程的结论.

3. 证明 (3.2.4) 关于指数随机变量的结论.

4. 设 $W(t) : \mathbb{R}^+ \mapsto \mathbb{R}$ 是一个标准 Brownian 运动. 计算 $W(t) - W(s)\,(t > s \geqslant 0)$ 所有的矩.

5. 计算第 4 题中 $W(t) : \mathbb{R}^+ \mapsto \mathbb{R}$ 是一个 d 维 Brownian 运动时, $W(t) - W(s)\,(t > s \geqslant 0)$ 所有的矩.

6. 设 $W(t)$ 是一个一维标准 Brownian 运动. 对于 $\Delta W(t) = W(t + \Delta t) - W(t)$ 有

$$\mathbb{E}\left(\left|\frac{\Delta W(t)}{\Delta t}\right|\right) = \sqrt{\frac{2}{\pi \Delta t}},$$

推出 Brownian 运动处处不可导.

7. 设 $W(t)$ 在 $\mathbb{R}^d$ 上是一个一维标准 Brownian 运动, 并且设 $\mathcal{F}_t$ 为 $W(t)$ 产生的过滤. 证明 $W(t)$ 以及 $W(t) \otimes W(t) - It$ 都是 $\mathcal{F}_t$ 鞅.

8. 说明在任意维情形下有类似定理 3.32 的结论.

9. 考虑 Brownian 桥 $BB(t) : [0, T] \mapsto \mathbb{R}$, 其定义如下:

$$BB(t) = W(t) - \frac{t}{T} W(T),$$

其中 $W(t)$ 是一个一维标准 Brownian 运动. 计算 Brownian 桥的均值和协方差.

第 4 章　常微分方程

4.1　格　　局

本章研究如下形式的常微分方程:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = h(z), \quad z(0) = z_0, \tag{4.1.1}$$

其中 $h: \mathcal{Z} \to \mathbb{R}^d$, 当 $\mathcal{Z} = \mathbb{T}^d, \mathbb{R}^d$ 或 $\mathbb{R}^l \oplus \mathbb{T}^{d-l}$ 时是其中比较典型的情况. 将在后面的章节中讨论 $z = (x^{\mathrm{T}}, y^{\mathrm{T}})^{\mathrm{T}}$ 的情形, 其中 $x \in \mathcal{X}, y \in \mathcal{Y}$. 如果 $\mathcal{Z} = \mathbb{T}^d$(或 $\mathcal{Z} = \mathbb{R}^d$), 则 $\mathcal{X} = \mathbb{T}^l$(或 $\mathcal{X} = \mathbb{R}^l$) 以及 $\mathcal{Y} = \mathbb{T}^{d-l}$(或 $\mathcal{Y} = \mathbb{R}^{d-l}$). 如果 $\mathcal{Z} = \mathbb{R}^l \oplus \mathbb{T}^{d-l}$, 则 $\mathcal{X} = \mathbb{R}^l$ 且 $\mathcal{Y} = \mathbb{T}^{d-l}$.

当在胞体上考虑方程 (4.1.1) 时, 该方程是一种简略形式, 其中 h 在 z(或 y) 上具有周期性, 并且在各方向上的周期为 1. 有时在胞体上考虑微分算子, 这时算子会自动具有周期边界条件.

在 4.2 节中列举了常微分方程解的存在唯一性理论; 生成子思想 (Liouville 等式) 以及特征线法将在 4.3 节中介绍; 在 4.4 节里讨论了常微分方程的遍历性; 对本章中出现的一些结论的引申连同参考书目见 4.5 节.

4.2　存在性和唯一性

当对所有在 $\mathcal{Z}$ 上的初值 z_0 以及 $t \in \mathbb{R}$, 初值问题 (4.1.1) 存在唯一解 $z(t)$ 时, 记

$$z(t) = \varphi^t(z_0),$$

这样 $\varphi^t: \mathcal{Z} \to \mathcal{Z}$ 就是解算子, 并且形成单参数算子群, 即

$$\varphi^{t+s} = \varphi^t \circ \varphi^s, \quad \forall t, s \in \mathbb{R} \text{ 且 } \varphi^0 = \mathcal{I},$$

其中 $\mathcal{I}: \mathcal{Z} \to \mathcal{Z}$ 为恒等算子, φ^t 的逆为 φ^{-t}. 通过如下定义可能会在集合上应用 φ^t:

$$\varphi^t(A) = \bigcup_{z \in A} \varphi^t(z).$$

一旦定义了 φ^t, 通常就将初值条件为 z 的方程 (4.1.1) 的解 $\varphi^t(z)$ 简记为 φ^t ①.

① 在这种情况下, 不应将 z 与 $z(t)$ 混淆.

实际上, 在不同情况下可以证明方程 (4.1.1) 的解是存在且唯一的. 为了便于说明问题, 在这里作最简单的假设, 即 h 在 $\mathcal{Z}$ 上是 Lipschitz 连续的. 若解在有限时间内能够阻止数据的放大, 这个条件就可以被减弱, 可以用 Lyapunov 函数证明这种先验有界性, 见 4.3 节.

定义 4.1 如果

$$|f(x)-f(y)| \leqslant L|x-y|, \quad \forall x, y \in \mathcal{Z},$$

则函数 $f: \mathcal{Z} \rightarrow \mathbb{R}^d$ 在 $\mathcal{Z}$ 上是 Lipschitz 连续的, Lipschitz 常数为 L.

注意 Lipschitz 函数也具有线性边界,

$$|f(x)| \leqslant |f(0)|+L(x).$$

定理 4.2 如果在 $\mathcal{Z}$ 上 h 是 Lipschitz 函数, 那么对所有 $z_0 \in \mathcal{Z}$, 常微分方程 (4.1.1) 有唯一解, 并且当 $z(t) \in C^1(\mathbb{R}, \mathcal{Z})$ 时有

$$|\varphi^t(x)-\varphi^t(y)| \leqslant e^{Lt}|x-y|, \quad \forall x, y \in \mathcal{Z}, t \geqslant 0. \tag{4.2.1}$$

证明 可由 Picard 迭代得出解的存在唯一性. 为了获取有界性 (4.2.1), 注意到

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d} t}|x-y|^2=\langle h(x)-h(y), x-y\rangle .$$

使用 h 的 Lipschitz 特性可得

$$\langle h(x)-h(y), x-y\rangle \leqslant L|x-y|^2,$$

从而对 $e=|x-y|^2$ 有

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d} t}|e| \leqslant L|e|.$$

对这个微分不等式进行积分 (或利用 Gronwall 引理 4.4) 得出所求结果.

设 f 只在 $\mathcal{Z}$ 的紧子集上是 Lipschitz 的, 通常对于任意 $z \in \mathcal{Z}$, 可在一些包含 0 的区间 $t \in [-a, b]$ 上定义 $\varphi^t(z)$, 一般地, 该区间依赖于 z. 有了这些假设就可得到如下定义:

定义 4.3 如果 $t \in \mathbb{R}(t \in [0, \infty)$ 或 $\in(-\infty, 0])$, 对所有的 $z \in \mathcal{A}$ 定义 $\varphi^t(z)$, 并且对于所有 $t \in \mathbb{R}\,(t \in (0, \infty]$ 或 $t \in[-\infty, 0))$, $\varphi^t(\mathcal{A}) \equiv \mathcal{A}($ 或者 $\varphi^t(\mathcal{A}) \subseteq \mathcal{A})$, 则称集合 $\mathcal{A}$ 是不变的 (向前不变、向后不变).

研究进化方程的一个很重要的工具是 Gronwall 引理. 它是一个指数解的不等式形式, 像前面定理说明的那样, 通过解线性常数微分方程可以得到该指数解.

引理 4.4 (Gronwall)

(i)(微分形式) 设 $\eta(t) \in C^1([0,T];\mathbb{R}^+)$ 满足下面的微分不等式:

$$\frac{\mathrm{d}\eta(t)}{\mathrm{d}t} \leqslant a\eta(t) + \psi(t), \quad \eta(0) = \eta, \tag{4.2.2}$$

其中 $a \in \mathbb{R}$ 且 $\psi(t) \in L^1([0,T];\mathbb{R}^+)$, 则对于所有 $t \in [0,T]$ 有

$$\eta(t) \leqslant \exp(at)\left(\eta + \int_0^t \exp(-as)\psi(s)\mathrm{d}s\right); \tag{4.2.3}$$

(ii)(积分形式) 假定对于某些正常数 $a, b, \xi(t) \in C([0,T];\mathbb{R}^+)$ 满足积分不等式①

$$\xi(t) \leqslant a\int_0^t \xi(s)\mathrm{d}s + b,$$

则

$$\xi(t) \leqslant b\exp(at), \quad t \in [0,T].$$

证明 (i) 用 $\exp(-at)$ 乘以式 (4.2.2) 得到

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\eta(t)\right)\exp(-at) \leqslant (\psi(t) + a\eta(t))\exp(-at),$$

从而

$$\frac{\mathrm{d}}{\mathrm{d}t}(\eta(t)\exp(-at)) \leqslant \exp(-at)\psi(t).$$

从 0 到 t 上积分上述不等式并乘以 $\exp(at)$ 就可得 (4.2.3).

(ii) 定义 $\eta(t) = \int_0^t \xi(s)\mathrm{d}s$, 则 $\eta(t)$ 满足不等式

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} \leqslant a\eta + b.$$

由于 $\eta(0) = 0$, 运用 (i) 可得

$$\eta(t) \leqslant \frac{b}{a}(\exp(at) - 1);$$

从而

$$\xi(t) \leqslant a\eta(t) + b \leqslant b\exp(at). \qquad \square$$

---

① 这种假设可能会被减弱.

4.3 生　成　子

理解 $t \mapsto z(t)$ 的函数如何随时间变化是很重要的. 通过使用如下**生成子**$\mathcal{L}$:

$$\mathcal{L}v = h(z) \cdot \nabla v(z) \tag{4.3.1}$$

可以达到这个目的.

如果 $z(t)$ 是 (4.1.1) 的解且 $V \in C^1(\mathcal{Z}, \mathbb{R})$, 则

$$\frac{\mathrm{d}}{\mathrm{d}t}(V(z(t))) = \nabla V(z(t)) \cdot \frac{\mathrm{d}z}{\mathrm{d}t}(t) = \nabla V(z(t)) \cdot h(z(t)) = \mathcal{L}V(z(t)). \tag{4.3.2}$$

在 $\mathbb{R}^n$ 上取值, 则生成子易成为函数. 有时对于 $V(z)$, 所有的 $z \in \mathcal{Z}$, $\mathcal{L}V(z)$ 有上界. 如果 V 也是正的, 则可采用微分不等式 (如 Gronwall 引理) 来获得 $V(z(t))$ 的边界, 从而得到 $z(t)$ 的边界, 那么 $V(z(t))$ 就是一个 Lyapunov 函数.

例 4.5　考虑方程

$$\frac{\mathrm{d}z}{\mathrm{d}t} = z - z^3. \tag{4.3.3}$$

如果 $V(z) = \frac{1}{2}z^2$, 则

$$\mathcal{L}V(z) = z^2 - z^4 \leqslant 1 - 2V(z),$$

从而

$$\frac{\mathrm{d}}{\mathrm{d}t}(V(z(t))) \leqslant 1 - 2V(z(t)).$$

由 Gronwall 引理得

$$V(z(t)) \leqslant \mathrm{e}^{-2t}V(z(0)) + \frac{1}{2}(1 - \mathrm{e}^{-2t}).$$

因此, 存在 Lyapunov 函数 $V(z) = \frac{1}{2}|z|^2$, 从而得到 (4.3.3) 的解的先验边界

$$|z(t)^2 \leqslant 1 + \mathrm{e}^{-2t}|z(0)|^2.$$

□

同样由下式给出的 L^2 伴随算子 $\mathcal{L}^*$ 也很重要:

$$\mathcal{L}^*v = -\nabla \cdot (hv), \tag{4.3.4}$$

其中定义的 $\mathcal{L}$ 和 $\mathcal{L}^*$ 是仅适用于 $C^1(\mathcal{Z})$ 中的函数. 至于将定义范围分别延伸到 $L^\infty(\mathcal{Z})$ 和 $L^1(\mathcal{Z})$ 上的函数时, 它也是非常有用的, 这将在后面进行讨论.

在理解式 (4.1.1) 的解集时 (其参数由初值决定并可能含有概率测度), $\mathcal{L}$ 和 $\mathcal{L}^*$ 起到关键作用, 现说明如下：设 v 是 Cauchy 问题 (见第 7 章) 的解,

$$\frac{\partial v}{\partial t} = \mathcal{L}v, \quad (z,t) \in \mathcal{Z} \times (0,\infty), \tag{4.3.5a}$$

$$v(z,0) = \phi(z), \quad z \in \mathcal{Z}. \tag{4.3.5b}$$

有时称之为**向后方程**. 用 $v(z,t) = (\mathrm{e}^{\mathcal{L}t}\phi)(z)$ 表示式 (4.3.5) 的解, 其通常被看成是时变线性算子方程的解所构成的半群, 见 7.5 节. 通过线性偏微分方程的古典解, 使得函数 $v(z,t)$ 充分光滑, 以至于对 $(z,t) \in \mathcal{Z} \times (0,\infty)$ 满足偏微分方程, 并且方程在 $\mathcal{Z} \times \mathbb{R}^+$ 上是连续的, 见第 7 章.

定理 4.6 假定 (4.1.1) 的解在 $\mathcal{Z}$ 上生成一个单参数群, 并对于所有 $t \in \mathbb{R}$, 使得 $\varphi^t(\mathcal{Z}) = \mathcal{Z}$. 再设 ϕ 充分光滑, 使得 (4.3.5) 有古典解, 则古典解可由下式得出①:

$$v(z,t) = \phi(\varphi^t(z)), \quad \forall t \in \mathbb{R}^+, z \in \mathcal{Z}. \tag{4.3.6}$$

证明 注意到 (4.3.6) 满足初值条件 $v(z,0) = \phi(\varphi^0(z)) = \phi(z)$. 由 φ^t 的群性质可得 $v(z,t)$ 满足 $v(\varphi^{-t}(z),t) = \phi(z), \forall t \in \mathbb{R}^+, z \in \mathcal{Z}$, 对 t 求导, 根据 (4.3.6) 得

$$\frac{\mathrm{d}}{\mathrm{d}t}(v(\varphi^{-t}(z),t)) = 0,$$

所以

$$\frac{\partial v\left(\varphi^{-t}(z),t\right)}{\partial t} + \left\langle \nabla v\left(\varphi^{-t}(z),t\right), \frac{\mathrm{d}}{\mathrm{d}t}\left(\varphi^{-t}(z)\right)\right\rangle = 0.$$

由于 $\varphi^{-t}(z)$ 是 (4.1.1) 向后时间解, 因此, 满足

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\varphi^{-t}(z)\right) = -h\left(\varphi^{-t}(z)\right),$$

则

$$\frac{\partial v\left(\varphi^{-t}(z),t\right)}{\partial t} + \left\langle \nabla v\left(\varphi^{-t}(z),t\right), -h\left(\varphi^{-t}(z)\right)\right\rangle = 0, \quad \forall t \in \mathbb{R}^+, z \in \mathcal{Z},$$

等价于

$$\frac{\partial v(z,t)}{\partial t} + \langle \nabla v(z,t), -h(z)\rangle = 0, \quad \forall t \in \mathbb{R}^+, z \in \mathcal{Z}.$$

表明 (4.3.6) 为线性偏微分方程 (4.3.5) 的解.

注 4.7 式 (4.3.6) 用**特征线法**表示偏微分方程 (4.3.5) 的解, 显然表明非线性常微分方程 (4.1.1) 的解集 (对所有 $z_0 \in \mathcal{Z}$) 可由线性偏微分方程的解表示. 相反,

① 注意: 此处 z 是偏微分方程中的独立空间变量, 其作为一个群 ϕ^t 的初始条件出现在偏微分方程解的公式中. z 的这种用法不应当与函数 $z(t)$ 混淆, $z(t)$ 是 (4.1.1) 的解, 其初始条件为 z_0.

也可以通过解常微分方程的非线性系统来求解具有初值问题的线性进化偏微分方程, 该方程在空间和时间上是一阶的 (运输方程). □

通过设

$$(\mathrm{e}^{\mathcal{L}t}\phi)(z)=\phi(\varphi^t(z)),\quad \forall t\in\mathbb{R}^+,z\in\mathcal{Z}, \tag{4.3.7}$$

可将算子 $\mathrm{e}^{\mathcal{L}t}$ 的定义延伸到任意函数 $\phi\in L^\infty(\mathcal{Z})$, 其中算子 $\mathrm{e}^{\mathcal{L}t}$ 将 $L^\infty(\mathcal{Z})$ 映射到它本身.

接下来定义生成子 $\mathcal{L}$, 设函数 ϕ 在 $L^\infty(\mathcal{Z})$ 上的极限

$$\mathcal{L}\phi=\lim_{t\to 0}\frac{\mathrm{e}^{\mathcal{L}t}\phi-\phi}{t} \tag{4.3.8}$$

存在, 令 $\mathcal{D}(\mathcal{L})$ 表示这种函数 ϕ 的集合. 上述这种极限在 $\mathcal{D}(\mathcal{L})$ 上定义了 $\mathcal{L}$, 它和 $C^1(\mathcal{Z})$ 上由 (4.3.1) 定义的 $\mathcal{L}$ 是一致的, 然而, 空间 $C^1(\mathcal{Z})$ 不是 $L^\infty(\mathcal{Z})$ 的一个稠密子集.

下面在 z_0 上设置概率测度, 使得 (4.1.1) 的解 $z(t)$ 是一个随机变量, 研究将会得到怎样的结果. 为此, 考虑 (4.3.5) 的伴随矩阵, 即 **Liouville 方程**:

$$\frac{\partial\rho}{\partial t}=\mathcal{L}^*\rho,\quad (z,t)\in\mathcal{Z}\times(0,\infty), \tag{4.3.9a}$$

$$\rho(z,0)=\rho_0(z),\quad z\in\mathcal{Z}. \tag{4.3.9b}$$

有时也称之为**向前方程**. 使用半群记号, 可用 $\rho(z,t)=(\mathrm{e}^{\mathcal{L}^*t}\rho_0)(z)$ 来表示解. 因为 $\mathcal{L}^*$ 是 $\mathcal{L}$ 的伴随算子, 所以 $\mathrm{e}^{\mathcal{L}^*t}$ 是 $\mathrm{e}^{\mathcal{L}t}$ 的伴随 (见 7.5 节). 令 $\mathbb{E}$ 为初值的期望, 其分布服从于 $\mathcal{Z}$ 上的随机变量, 概率密度为 $\rho_0(z)$ ①, 即

$$\mathbb{E}f:=\int_{\mathcal{Z}}f(z)\rho_0(z)\mathrm{d}z.$$

定理 4.8 假定 (4.1.1) 的解在 $\mathcal{Z}$ 上生成单参数群, 并使得对所有 $t\in\mathbb{R}$ 有 $\varphi^t(\mathcal{Z})=\mathcal{Z}$. 再设 ϕ 充分光滑, 使得 (4.3.5) 有古典解. 最后设 (4.1.1) 的初值是一个在 $\mathcal{Z}$ 上的随机变量, 服从概率密度为 $\rho_0(z)$ 的分布, 并且充分光滑, 使得 (4.3.9) 有古典解, 则 $z(t)$ 是 $\mathcal{Z}$ 上的一个随机变量, 服从概率密度为 $\rho(z,t)$ 的分布且满足 (4.3.9).

证明 由定理 4.6 有

$$\begin{aligned}\mathbb{E}\left(\phi\left(z(t)\right)\right)&=\int_{\mathcal{Z}}\phi\left(\varphi^t(z)\right)\rho_0(z)\mathrm{d}z=\int_{\mathcal{Z}}v\left(z,t\right)\rho_0(z)\mathrm{d}z\\&=\int_{\mathcal{Z}}\left(\mathrm{e}^{\mathcal{L}t}\phi\right)(z)\rho_0(z)\mathrm{d}z=\int_{\mathcal{Z}}\left(\mathrm{e}^{\mathcal{L}^*t}\rho_0\right)(z)\phi(z)\mathrm{d}z.\end{aligned}$$

① 再一次注意, 此处 z 是一个的独立空间变量, 不应当将其与函数 $z(t)$ 混淆, 其中 $z(t)$ 是 (4.1.1) 的解, 其初始条件为 z_0.

另外, 如果 $\rho(z,t)$ 是 $z(t)$ 的密度函数, 则

$$\mathbb{E}(\phi(z(t))) = \int_{\mathcal{Z}} \rho(z,t)\,\phi(z)\mathrm{d}z.$$

在时间 t 上令两个表达式相等, 利用 ϕ 的任意性和在 $L^2(\mathcal{Z})$ 上所有 ϕ 满足等式, 则在 $L^2(\mathcal{Z})$ 上可得

$$\rho(z,t) = (\mathrm{e}^{\mathcal{L}^* t}\rho_0)(z),$$

因此, 由光滑性假设可知密度 $\rho(z,t)$ 满足方程 (4.3.9).

注 4.9 对 $\mathrm{e}^{\mathcal{L}^* t}$ 的定义域进行延伸, 通过定义算子 $\mathrm{e}^{\mathcal{L}t}: L^\infty(\mathcal{Z}) \to L^\infty(\mathcal{Z})$(见 (4.3.7)) 的 L^2 伴随算子, 将 $L^1(\mathcal{Z})$ 映射到本身, 结果得到算子 $\mathrm{e}^{\mathcal{L}^* t}$ 在 $L^1(\mathcal{Z})$ 上是连续的.

由如下方式定义生成子 $\mathcal{L}^*$ 的 L^2 伴随: 设函数 ρ 在 $L^1(\mathcal{Z})$ 上的极限

$$\mathcal{L}^*\rho = \lim_{t\to 0} \frac{\mathrm{e}^{\mathcal{L}^* t}\rho - \rho}{t} \tag{4.3.10}$$

存在, 令 $\mathcal{D}(\mathcal{L}^*)$ 表示这种函数 ρ 的集合. 它和 $C^1(\mathcal{Z})$ 上由 (4.3.4) 定义的 $\mathcal{L}^*$ 是一致的, 并且 $C^1(\mathcal{Z})$ 是 $L^1(\mathcal{Z})$ 的稠密子集. □

4.4 遍 历 性

简单地说, 遍历性是讨论动力系统一种测度不变量的存在唯一性, 即不变测度的存在唯一性. 这种带有启发式的考虑遍历性的思考方式, 有利于在本章及后面两章中将那些稍有区别的遍历性表达式统一起来. 尤其在考虑非随机问题时, 采用这种非常有技术性的思考方式有助于理解第二部分的内容, 此时遍历性在许多扰动展开中起到关键作用. 我们会将这种思考方式与许多其他的相关概念联系起来, 尤其是在关于不变测度下时间平均的收敛性、生成子的一维零空间及其在适度空间上的伴随算子等.

在本章中, 考虑测度空间 $(\mathcal{Z}, \mathcal{A}, \mu)$, 其中 μ 为 $\mathcal{Z}$ 上的一个测度, 并且 $\mathcal{A}$ 为 $\mathcal{Z}$ 的 μ 可测子集的一个 σ 代数, 设 φ^t 为 (4.1.1) 的解算子.

定义 4.10 如果

$$\mu(\varphi^t(A)) = \mu(A), \quad \forall A \in \mathcal{A}, t > 0,$$

则称测度 μ 是不变的.

回想如果对于所有 $t \in \mathbb{R}$, 并且有

$$\varphi^t(A) = A,$$

则称集合 $A \in \mathcal{A}$ 在 φ^t 下是不变的.

定义 4.11　给定不变测度 μ, 如果对于 φ^t 的每个不变集合 A 有 $\mu(A)=0$ 或 $\mu(A)=1$, 则称常微分方程 (4.1.1) 具有遍历性.

注意到遍历性定义与所研究问题的测度空间有关, 这样就出现了两个常见问题: (i) 如何确定一个测度 μ 是不变测度? (ii) 给定一个不变测度 μ, 怎样检验其遍历性? 通常来说, 对于常微分方程这些都是很难的问题.

在本节剩余部分, 作简单的假设, 即 $\mathcal{Z}=\mathbb{T}^d$ 且由 (4.3.1) 给出算子 $\mathcal{L}$ 和 $\mathcal{L}^*$, 同时 (4.3.4) 具有周期边界条件. 接下来的两个定理依次讨论前面提出的两个问题, 尤其是在周期条件下. 在下文中, 区分 L^p 空间中的 Lebesgue 测度以及不变测度 μ(通常其是不同的) 是非常重要的.

定理 4.12　设 μ 为不变概率测度,

(i) 如果 $\phi \in L^1(\mathbb{T}^d;\mu)$, 则对于 μ a.e. $z \in \mathbb{T}^d$ 有

$$\phi^*(z)=\lim_{T\to\infty}\frac{1}{T}\int_0^T \phi(\varphi^t(z))\mathrm{d}t;$$

(ii) 带有密度函数 $\rho^\infty \in L^1(\mathbb{T}^d;\mathrm{Leb})$ 的测度 μ 关于 Lebesgue 测度是不变的充要条件为

$$\mathrm{e}^{\mathcal{L}^*t}\rho^\infty=\rho^\infty, \quad \rho^\infty \in L^1(\mathbb{T}^d;\mathrm{Leb}), \forall t \geqslant 0;$$

(iii) 带有密度函数 $\rho^\infty \in C^1(\mathbb{T}^d)$ 的测度 μ 关于 Lebesgue 测度是不变的充要条件为

$$\mathcal{L}^*\rho^\infty=0, \quad \rho^\infty \in C^1(\mathbb{T}^d).$$

注意: (iii) 具有很好的特性, 即只需要验证 ρ^∞ 是连续可微的, 那么就可使用式 (4.3.4) 中的微分算子 $\mathcal{L}^*$. 这样做实质上是由 $L^1(\mathbb{T}^d;\mathrm{Leb})$ 上算子 $\mathrm{e}^{\mathcal{L}^*t}$ 的连续性以及 $L^1(\mathbb{T}^d;\mathrm{Leb})$ 中 $C^1(\mathbb{T}^d)$ 的稠密性得出来的. 然而, 当验证遍历性时, 不可能将 $\mathcal{L}$ 作为微分算子使用. 在下面的定理中, 必须对所有 $\mathcal{L}$ 定义域上的函数验证条件 (iii), 而不仅仅是只要使 (4.3.1) 成立就可以了.

定理 4.13　设 φ^t 有不变概率测度 μ,

(i) 如果 φ^t 具有遍历性且 $\phi \in L^1(\mathbb{T}^d;\mu)$, 则

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \phi(\varphi^t(z))\mathrm{d}t=\int_{\mathbb{T}^d}\varphi(z)\mu(\mathrm{d}z), \tag{4.4.1}$$

其中 μ a.e. $z \in \mathbb{T}^d$;

(ii) φ^t 具有遍历性的充要条件是方程

$$\mathrm{e}^{\mathcal{L}t}g=g, \quad g \in L^\infty(\mathbb{T}^d), \forall t \geqslant 0$$

有唯一常数解 μ a.e.;

(iii) φ^t 具有遍历性的充要条件是方程

$$\mathcal{L}g = 0, \quad g \in \mathcal{D}(\mathcal{L})$$

有唯一常数解 μ a.e..

注意前面定理概括的第一条在遍历性中是一个很重要的性质, 即不变测度的平均可以代替时间平均. 在第二部分中, 大部分对常微分方程遍历性的推导会使用这个观点; 当研究随机微分方程中类似的问题时也会使用时间平均表达式. 还注意到与前面两个定理第一项相比, 当有不变测度时, 时间平均总是收敛的, 但遍历性的一个重要特征是时间平均收敛于一个值, 该值独立于初始条件.

对于遍历性系统, 动态轨迹掠过整个相位空间, 其独立于初始条件且在任意子集上停留相等时间. 令 ϕ 为 I_A(I_A 为 Borel 集, $A \subseteq \mathbb{T}^d$) 的指标函数, 由定理 4.13 推出下面公式给出的测度 μ:

$$\mu(A) = \lim_{T\to\infty} \frac{1}{T}\int_0^T I_A(z(t))\mathrm{d}t, \tag{4.4.2}$$

则不变测度 μ 度量了时间比重, 表示了典型轨迹停留在给出的 $\mathbb{T}^d$ 子集里的时间.

例 4.14 考虑方程

$$\frac{\mathrm{d}z}{\mathrm{d}t} = b(z),$$

其中 $b \in C^\infty(\mathbb{T}, \mathbb{R})$ 且

$$\inf_{z\in\mathbb{T}} b(z) > 0.$$

对每个固定的 $t \in \mathbb{R}$, 记 $z(t)$ 为环面 (圆环)$\mathbb{T}$ 的元, 此处

$$\mathcal{L}\phi = b\frac{\partial \phi}{\partial z}$$

具有周期边界条件. 定理 4.12(iii) 表明方程有不变测度 μ, 其概率密度为

$$\rho^\infty(z) = \frac{C}{b(z)},$$

其中选择 C 以便于 ρ^∞ 在 $\mathbb{T}$ 上积分等于 1. 很自然地, 还需要确定此不变测度方程是否具有遍历性.

首先考虑 $b(z) = \omega > 0$ 是一个常数的情形, 则 μ 仅仅是 Lebesgue 测度. 注意圆环 $\mathbb{T} = \mathbb{S}$, 则

$$\varphi^\tau(z) = z + \omega\tau \ (\mathrm{mod}\ 1).$$

显然, 对任意不变 τ, 旋转 $\varphi^\tau : \mathbb{S} \to \mathbb{S}$ 是一个遍历性映射, 使得 $\omega\tau$ 是无理数, 即对于任意满足 $\varphi^\tau(A) = A$ 的集合 A 有 Lebesgue 测度 (因而也是 μ 测度) 且为 0 或 1.

考虑这样的 τ, 任意常微分方程的不变子集, 尤其当 $t=\tau$ 时, 必须满足 $\varphi^t(A)=A$, 则其有 Lebesgue 测度 (因而也是 μ 测度) 其为 0 或 1, 因而常微分方程在 $\mathbb{T}$ 上具有遍历性.

对于 $b(z)$ 不是常数的情形, 可能会重新标定时间以使向量场为常数, 由于 b 在 $\mathbb{S}$ 上是光滑的且严格为正, 这样处理是有可能的. 既然 μ 关于 Lebesgue 测度是绝对连续的, 对于重新标定的方程就暗示了初始方程也具有相同的性质. □

例 4.15　考虑方程 (4.1.1), 该方程被看成是 $\mathbb{T}^d$ 上的动力系统, 其中向量场 $h(z)$ 是无散度向量值, 即 $\nabla\cdot h(z)=0$, 则由 (4.3.1), (4.3.4) 给出的 $\mathcal{L}$ 和 $\mathcal{L}^*$ 满足

$$\mathcal{L}^*v=-\nabla\cdot(hv)=-(\nabla\cdot h)v-h\cdot\nabla v=-h\cdot\nabla v=-\mathcal{L}v.$$

两个算子都具有周期边界条件. 因此, 在 $\mathbb{T}^d$ 上方程 $\mathcal{L}^*\rho^\infty=0$ 有解, 为 $\rho^\infty=1$, 从而对于常微分方程, Lebesgue 测度是不变的. 然而, 为了证明遍历性, 有必要说明方程 $\mathcal{L}g=0$ 在 $\mathcal{D}(\mathcal{L})\subset L^\infty$ 上有唯一解, 至于什么时候有唯一解就无关紧要了. 在第 14 章中, 当研究双曲方程时再讨论这个问题. □

到目前为止, 由于其主要集中在 $\mathcal{Z}$ 上且遍历性测度密度为 ρ^∞(关于 Lebesgue 测度) 的情况, 从而在应用过程中会受到某些限制. 这里有许多例子说明了遍历性的特性, 但是基础测度 μ 关于 Lebesgue 测度不是绝对连续的. 尽管如此, 这些问题说明一个重要性质, 即时间平均收敛于一个值, 该值独立于初值条件. (4.1.1) 中的常微分方程就是一个简单的例子, 该方程有唯一全局平衡解. 如果 μ 被选为平衡点上的 Dirac 质量, 则 (4.4.1) 成立. 现给出两个例子进一步说明这类性质.

例 4.16　考虑 Lorenz 方程

$$\begin{cases}\dfrac{\mathrm{d}y_1}{\mathrm{d}t}=10(y_2-y_1),\\[2mm]\dfrac{\mathrm{d}y_2}{\mathrm{d}t}=28y_1-y_2-y_1y_3,\\[2mm]\dfrac{\mathrm{d}y_3}{\mathrm{d}t}=y_1y_2-\frac{8}{3}y_3.\end{cases}\tag{4.4.3}$$

这些方程具有遍历性, 其不变测度在零空间及在 2~3 的分形维数上成立. 在这样的问题里, SRB 测度的概念与前面例子中不变测度的概念一样. 解的时间平均收敛于采用 SRB 测度[①]所得到的平均值, 因而 (4.4.1) 普遍成立.

图 4.1 和图 4.2 说明了这些思想. 图 4.1 表示 Lorenz 方程的吸引子, 其为不变测度成立的集合. 分形特征在集合的带状结构上是明显的, 并且被投影在 y 的前两个分量上. 图 4.2 表示由 Lorenz 方程的第二分式得到的经验测度 (矩形图), 开始于吸引子的初值. □

① 该测度来源于 Sinai, Ruelle 以及 Bowen.

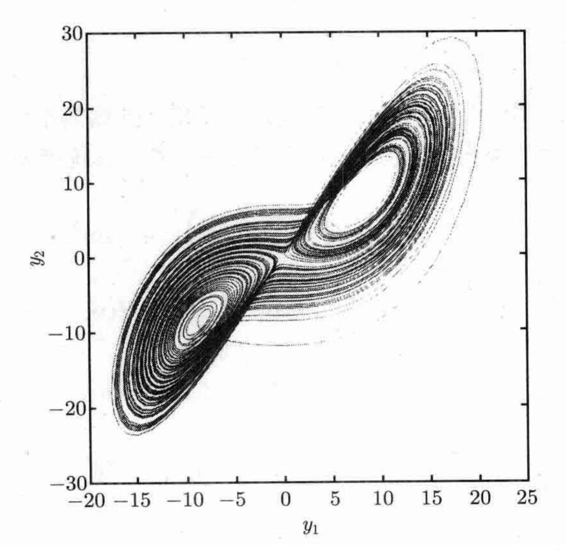


图 4.1 Lorenz 方程 (4.4.3) 的吸引子在 (y_1, y_2) 上的投影

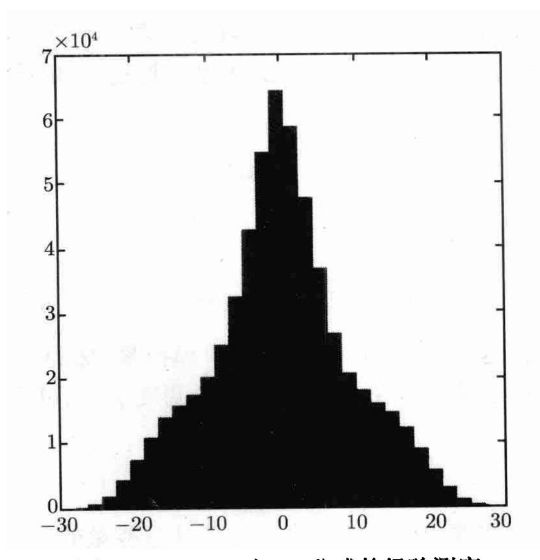


图 4.2 (4.4.3) 中 y_2 分式的经验测度

例 4.17 考虑谐振荡

$$\frac{\mathrm{d}^2\eta}{\mathrm{d}t^2} + \omega\eta = 0.$$

一阶系统可以记为

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = v,$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\omega\eta.$$

该方程有无散向量场, 因此, Lebesgue 测度是不变的 (虽然其不可规范化, 并且也不一定是概率测度), 而且微分算子 $\mathcal{L}$, $\mathcal{L}^*$ 满足 $\mathcal{L} = -\mathcal{L}^*$, 系统能量守恒,

$$E = \frac{1}{2}v^2 + \frac{1}{2}\omega\eta^2 = \frac{1}{2}\left(\frac{\mathrm{d}\eta}{\mathrm{d}t}\right)^2 + \frac{1}{2}\omega\eta^2.$$

直接求解 (见练习第 2 题 (c)) 表明 Hamiltonian 函数 $H(v,\eta)$ 的任意光滑函数在 $\mathcal{L}$ 的零空间上, 因此, 根据定理 4.13, 该方程具有遍历性.

然而, 方程满足能量壳 $\mathcal{E}$ 上的一种遍历性, $\mathcal{E}$ 由 $H(v,\eta) := E = \text{const.}$ 时给出, 这些可由方程

$$\eta(t) = A\sin\left(\omega^{1/2}t\right) + B\cos\left(\omega^{1/2}t\right)$$

和

$$\frac{\omega}{2}(A^2+B^2) = E$$

给出的解直观地进行理解. 当解在 $\mathcal{E}$ 上旋转就会生成一个经验测度 μ, 该测度量化了轨迹在 $\mathcal{E}$ 子集上掠过的时间, 结果表明此测度仅通过 E 依赖于 A 和 B. 直接求解得

$$\eta(t)^2 = \frac{1}{2}\left(A^2+B^2\right) + AB\sin\left(2\omega^{1/2}t\right) + \frac{1}{2}\left(B^2-A^2\right)\cos\left(2\omega^{1/2}t\right),$$

因而

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \frac{1}{2}\omega\eta^2(t)\mathrm{d}t = \frac{1}{4}\omega\left(A^2+B^2\right) = \frac{1}{2}E.$$

在 Hamiltonian 问题里, 这是 (4.4.1) 的一个模拟量, 然而, 注意通过 E, 初值条件的记忆还存在. 假定系统的确在 $\mathcal{E}$ 上生成经验测度 μ, 则可推导出

$$\int_{\mathcal{E}} \frac{\omega}{2}\eta^2\mu(\mathrm{d}\eta,\mathrm{d}v) = \frac{1}{2}E.$$

实际上, 由 **Virial 定理**得出振荡的平均动能是总能量的一半. 该定理表明动能和势能平等地分享整个能量. □

4.5 讨论和参考

定理 4.2 的完整证明见文献 [15, sec.3] 或 [68]; [188, sec.7.6] 中讨论了常微分方程系统的生成子及其性质; 利用生成子研究 Lyapunov 函数并得到常微分方程的先验性估计[305]; 有时方程 (4.3.1) 定义的生成子 $\mathcal{L}$ 及其伴随 $\mathcal{L}^*$ 分别被称为

Koopman 生成子和 **Frobenius-Perron** 算子; Koopman 算子也称为**转移算子**. 为了研究古典动力系统遍历性性质, Koopman 于 1931 年引入这些概念[176]; Koopman 原著及其理论延伸见文献 [266, sec. II.5, sec. VII.4].

Liouville 方程是非平衡统计力学的基本方程, 见文献 [28]; 通过对 Liouville 方程进行适当的模式消元程序, 可得 Fokker-Planck 方程[269,ch.9], 该方程是随机动力学的基本方程, 在第 6 章中将研究该方程.

在很多偏微分方程的书籍中都介绍了特征线法, 包括文献 [98].

有关遍历性的详细论述见文献 [188], 特别地, 定理 4.12 及定理 4.13 分别由[188]的定理 7.3.1, 定理 7.7.1, 定理 7.8.2 和定理 4.2.4, 定理 7.7.2, 定理 7.8.3 得出; 文献 [260, 327] 对离散和连续时间系统的遍历性理论进行了更详细的讨论, 也可参见文献 [14, 17, 290]; 有关 SRB 测度及其相关参考书见文献 [338].

Lorenz 方程的遍历性见 Tucker 的著作文献 [318, 319] 或者文献 [298]; 有关动力系统尤其是常微分方程的随机特性的文献综述见文献 [224]; 在许多经典力学的书籍中, 都可以找到对 Virial 定理的证明.

4.6 练 习

1. 设 $\mathcal{Z} = \mathbb{T}^d$, 证明对于所有 $f \in C^1(\mathcal{Z}, \mathbb{R})$, 方程 (4.3.1) 中定义的 $\mathcal{L}$ 的 L^2 伴随就是方程 (4.3.4) 中定义的 $\mathcal{L}^*$.

2. 设 $H(p,q) : \mathbb{R}^{2d} \to \mathbb{R}$ 是一个光滑函数, 并且考虑 (Hamiltonian) 常微分方程

$$\dot{q} = \nabla_p H(p,q), \quad \dot{p} = -\nabla_q H(p,q). \tag{4.6.1}$$

(a) 写出生成子以及方程 (4.6.1) 生成子的伴随矩阵;

(b) 写出 Liouville 方程 (4.6.1);

(c) 证明每个 Hamiltonian 光滑函数 $H(p(t), q(t))$ 都是 Liouville 方程的解;

(d) Hamiltonian 系统具有遍历性吗?

3. 根据第 2 题, 考虑 (梯度) 常微分方程

$$\dot{q} = -\nabla_q V(q) \tag{4.6.2}$$

时 (a), (b) 和 (d), 其中 $V(q) : \mathbb{R}^d \to \mathbb{R}$ 是一个光滑函数.

第 5 章 Markov 链

5.1 格　局

本节介绍可数状态空间上的 Markov 链. 不失一般性, 设这个状态空间为 $\mathcal{I}$, 为正整数集 $\mathbb{N}$ 的一个子集①.

在 5.2 节中, 通过研究离散时间 Markov 链引入一个思想; 在 5.3 节中, 从离散时间体中构造一个连续时间 Markov 链的例子; 在 5.4 节中, 用前面的例子作为一般的定义, 研究在 $\mathcal{I}$ 上的连续时间 Markov 链; 在 5.5 节中, 使用上述概念, 讨论 Markov 链解的存在唯一性; 5.6 节对有限态 Markov 链的遍历性进行了概述; 有关本章结论的各种延伸及参考文献见 5.7 节.

5.2 离散时间 Markov 链

如果

$$\sum_j p_{ij} = 1, \quad \forall i \in \mathcal{I},$$

并且对于所有 $(i,j) \in \mathcal{I} \times \mathcal{I}, p_{ij} \in [0,1]$, 则称由 $p_{i,j}$ 所构成的矩阵 P 是随机矩阵.

定义 5.1 如果随机序列 $\{z_n\}_{n\geqslant 0}$ 是一个具有状态空间 $\mathcal{I}$ 的 Markov 过程, 并且其向量的分量由 $\mathcal{I}$ 的基数得到, 以及转移矩阵 P 满足

- z_0 的分布为 ρ_0;
- 当 $\mathbb{P}(z_n = i) > 0$ 时, 对于每个 $n \geqslant 0$ 有 $\mathbb{P}(z_{n+1} = j | z_n = i) = p_{ij}$,

则称随机序列 $\{z_n\}_{n\geqslant 0}$ 为一个初始分布为 ρ_0 的离散时间 Markov 链, 同时, 称转移矩阵 $\{p_{ij}\}_{i,j\in\mathcal{I}}$ 的元素为**转移概率**. 通过构造使得 P 为随机矩阵, 注意通过使用 Chapman-Kolmogorov 方程 (3.3.5) 并进行推导, 使得 $\left(P^k\right)_{ij} = \mathbb{P}\left(z_k = j | z_0 = i\right)$, 因此, P^k 也是随机矩阵且 $\mathbb{P}(z_n = j) = \left(\rho_0^{\mathrm{T}} P^n\right)_j$.

离散时间 Markov 链有独立于 n 且从 z_n 到 z_{n+1} 的转移概率, 因此, 按照第 3 章的结论, 在离散集合 $\mathcal{I}$ 上产生的 Markov 链是均匀的, 有时称其为离散时间均匀 Markov 链. 由于只考虑均匀 Markov 链及其推论, 因此, 在后面就简称为均匀化.

---

① 在本章中, 所有求和都是在空间 $\mathcal{I}$ 上进行的, 除非说明是在其相反的空间上求和. 但是, 在后面的章节中有时会发现, 找到成对的变址状态矢量空间, 将其作为整数的两个子集的乘积, 在该空间上求和是很方便的.

例 5.2 设 $\alpha, \beta \in [0,1]$, 考虑矩阵

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

显然, 这是一个随机矩阵, 并且是双态 Markov 链的转换矩阵. □

例 5.3(关于 $\mathbb{Z}$ 的对称随机行走) 关于 $\mathbb{Z}$ 的对称随机行走是有如下初始分布的 Markov 链:

$$\rho_0^i = \begin{cases} 1, & i=0, \\ 0, & \text{其他} \end{cases}$$

且转移密度为

$$P = \begin{cases} \dfrac{1}{2}, & |i-j|=1, \\ 0, & \text{其他}, \end{cases}$$

等价于

$$\mathbb{P}(z_0 = 0) = 1$$

且

$$\mathbb{P}(z_{n+1} = j | z_n = i) = \begin{cases} \dfrac{1}{2}, & j = i+1, \\ \dfrac{1}{2}, & j = i-1, \\ 0, & \text{其他}. \end{cases}$$ □

注意到随机矩阵满足

$$P\mathbf{1} = \mathbf{1}, \tag{5.2.1}$$

其中 $\mathbf{1}$ 表示单位元素为 $\mathbf{1} = (1, \cdots, 1)$ 的矢量. 加上 P 有正元素, 则暗含下面的基本等式:

$$|P|_\infty = 1. \tag{5.2.2}$$

5.3 连续时间 Markov 链

现在在 $\mathcal{I}$ 上研究连续时间 Markov 链.

定义 5.4 连续时间 Markov 链是一个具有状态空间 $E = \mathcal{I}$ 的 Markov 随机过程 $\{z(t)\}_{t\in\mathbb{R}^+}$.

为了引入连续时间 Markov 链的概念, 使用离散时间 Markov 链来构造连续时间 Markov 链. 对一些 $\lambda > 0$, 设独立同分布序列 $\{\tau_n\}_{n\geqslant 0}$ 的分布为 $\exp(\lambda)$, 并由

$t_{n+1}=t_n+\tau_n, t_0=0$ 定义 $\{t_n\}_{n\geqslant 0}$. 设 $\{z_n\}_{n\geqslant 0}$ 为 $\mathcal{I}$ 上的离散时间 Markov 链, 其独立于 $\{\tau_n\}_{n\geqslant 0}$, 并且设

$$z(t)=z_n, \quad t\in[t_n,t_{n+1}), \tag{5.3.1}$$

称之为跳跃链[①]. 注意 $z(t)$ 在 $\mathcal{I}$ 上取值且 $z(t)$ 是 càdlàg 过程. 由 $\{z_n\}_{n\geqslant 0}$ 的 Markovian 结构以及指数随机变量的性质得出 $z(t)$ 是 Markov 链.

非正式地, 记 $z(t)$ 为下面微分方程的解:

$$\frac{\mathrm{d}z}{\mathrm{d}t}=\delta(t-t_j)\left(k\left(z\left(t^-\right)\right)-z\left(t^-\right)\right), \tag{5.3.2}$$

其中 $k(z)$ 的分布为 $p(z,\cdot)$, $\left\{k\left(z\left(t_j^-\right)\right)\right\}_{j\geqslant 0}$ 相互独立且均独立于 $\{\tau_j\}$. 得到这种表示方式是因为在跳跃时间 t_j 上积分可得到

$$z\left(t_j^+\right)-z\left(t_j^-\right)=\lim_{\varepsilon\to 0}\int_{t_j-\varepsilon}^{t_j+\varepsilon}\delta(t-t_j)\left(k\left(z\left(t^-\right)\right)-z\left(t^-\right)\right)\mathrm{d}t=k\left(z\left(t_j^-\right)\right)-z\left(t_j^-\right),$$

从而 $z\left(t_j^+\right)=k\left(z\left(t_j^-\right)\right)$. 弄明白这个随机微分方程, 尤其是对于所有 $t>0$ 都有解, 这与理解跳跃时间 t_j 不能在有限时间上累积这个问题是紧密相关的. 在 5.4 节中再继续讨论此微分方程解的存在性问题.

现在为跳跃链找一个矩阵 $P(t)$ 的表达式, 其元素为

$$p_{ij}(t)=\mathbb{P}(z(t)=j|\,z(0)=i), \tag{5.3.3}$$

用 P 和 λ 以及输入跳跃链的参数表示 $P(t)$. 注意到, 由 (3.2.4) 指数随机变量的性质有

$$\mathbb{P}(k \text{ jumps in } [0,t])=\frac{\mathrm{e}^{-\lambda t}(\lambda t)^k}{k!},$$

则由于 $\left(P^k\right)_{ij}=\mathbb{P}(z_k=j|\,z_0=i)$ 有

$$p_{ij}(t)=\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\lambda t}(\lambda t)^k}{k!}\left(P^k\right)_{ij},$$

因而

$$P(t)=\mathrm{e}^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}P^k=\mathrm{e}^{-\lambda t(P-I)}=\mathrm{e}^{Lt},$$

其中 $L=\lambda(P-I)$. 称矩阵 L 为连续时间 Markov 链的**生成子**. 当状态空间是无限时, 理解矩阵 e^{Lt} 要求明白半群理论, 见 7.5 节. 注意到由 (5.2.2) 有

$$L=\lim_{t\to 0}\frac{P(t)-I}{t},$$

① 也可能构造跳跃链, 此处跳跃率 λ 是依赖于初始状态的.

其中极限是在算子范数的意义下得到, 该范数由 $\mathcal{I}$ 上的 l^∞ 范数导出①, 因而算子特性与定义 3.11 中是一致的. 在下节中说明将算子当作起点, 从该点定义连续时间 Markov 链.

5.4 生 成 子

下面概括生成子的特性, 然后将其看成连续时间 Markov 链的一般定义.

定义 5.5 如果

- $\sum_j l_{ij} = 0, \forall i \in \mathcal{I}$;
- $l_{ij} \in [0, \infty), \forall (i,j) \in \mathcal{I} \times \mathcal{I}, i \neq j$,

则称元素为 l_{ij} 的矩阵 $L : \mathcal{I} \to \mathcal{I}$ 是连续时间 Markov 链的生成子. 注意到由于 P 是随机矩阵, 则前面章节中构造的矩阵 L 满足前面定义中的所有准则.

生成子的概念与第 3 章给出的 Markov 过程生成子的定义有关. 通过第 3 章的讨论, 注意到 Markov 链的特征为其生成子, 定义暗含

$$l(i) := -l_{ii} \in [0, \infty), \quad \forall i \in \mathcal{I}. \tag{5.4.1}$$

给定生成子 L, 就可能找到一个离散时间 Markov 链以及一系列的独立指数随机变量, 结果, 一个相关的跳跃链就会生成连续时间 Markov 链路径, 其生成子为 L, 其中

$$l^* := \sup_i l(i) < \infty. \tag{5.4.2}$$

通过选择独立同分布指数随机变量, 然后对前面章节中给出的例子求逆, 就可以实现这些想法②. 现在举例说明. 如果 (5.4.2) 成立, 则对于任意 $\lambda > l^*$,

$$P = I + \lambda^{-1} L \tag{5.4.3}$$

是一个随机矩阵. 生成独立同分布序列 τ_j, 其中 τ_0 分布是一个指数随机变量 $\tau \sim \exp(\lambda)$. 现在用转移矩阵为 P 的 Markov 链生成 z_n, 并且通过 (5.3.1) 定义 $z(t)$. 如前验证,

$$\frac{P(t) - I}{t} \to L,$$

则由定义 5.5 中的生成子, 已经在 $\mathcal{I}$ 上构造了一个连续时间随机过程. 注意 $z(t)$ 在 $\mathcal{I}$ 上取值, 是 càdlàg 且可能由微分方程 (5.3.2) 描述.

现在用别的方式观察转移矩阵为 $P(t)$ 的连续时间 Markov 链和生成子 L 之间的联系. 考虑一个在状态空间 $\mathcal{I} \subseteq \{1, 2, \cdots\}$ 上取值的连续时间 Markov 链 $z(t)$ $(t \geqslant$

① 如果 $\mathcal{I}$ 是有限维的, 则可能会使用任意 l^∞ 范数.

② 当不满足 (5.4.1) 中的条件时, 必须选取依赖于初始状态的指数率.

0). 设 $p_{ij}(t)$ 为从状态 i 到 j 的转移概率, i,j 由 (5.3.3) 给出. 对于所有 $t,\Delta t \geqslant 0$, Markov 特性暗含

$$p_{ij}(t+\Delta t)=\sum_k p_{ik}(t)\,p_{kj}(\Delta t),$$

其中

$$\frac{p_{ij}(t+\Delta t)-p_{ij}(t)}{\Delta t}=\sum_k p_{ik}(t)\,l_{kj}(\Delta t),$$

$$l_{kj}(\Delta t)=\frac{1}{\Delta t}\times\begin{cases}p_{kj}(\Delta t), & k\neq j,\\ p_{jj}(\Delta t)-1, & k=j.\end{cases}\tag{5.4.4}$$

假设极限 $l_{kj}=\lim\limits_{\Delta t\to 0} l_{kj}(\Delta t)$ 存在, 则得

$$\frac{\mathrm{d}p_{ij}}{\mathrm{d}t}=\sum_k p_{ik}l_{kj}.\tag{5.4.5}$$

因为 $\sum\limits_j p_{ij}(\Delta t)=1$, 则得 $\sum\limits_j l_{ij}(\Delta t)=0$, 并且假设下面的极限存在:

$$\sum_j l_{ij}=0.\tag{5.4.6}$$

这暗含对于极限方程 (5.4.5) 有

$$\sum_j p_{ij}=1.$$

引入矩阵 $P(t)$, L 元素分别为 $p_{ij}(t), l_{ij}(i,j\in\mathcal{I})$. 由矩阵概念, 方程 (5.4.5) 为

$$\frac{\mathrm{d}P}{\mathrm{d}t}=PL,\quad P(0)=I.\tag{5.4.7}$$

(5.4.4) 表明, 通过下面的公式, L 可由 P 计算得到:

$$L=\lim_{\Delta t\to 0}\frac{1}{\Delta t}(P(\Delta t)-I)=\left.\frac{\mathrm{d}P}{\mathrm{d}t}\right|_{t=0}.$$

(L 和 P 之间的关系也可由后面章节中的逆构造得到). 由 (5.4.6) 可知在零空间上生成子有常数,

$$L\mathbf{1}=0.\tag{5.4.8}$$

当 P 是一个随机矩阵时得到

$$\left|\mathrm{e}^{Lt}\right|_\infty=1,\tag{5.4.9}$$

而且 p_{ij} 的非负性暗示 L 有非负非对角元, 则条件 (5.4.6) 暗示 L 的对角元也是非负的.

注意:

$$P(t)=\exp(Lt) \tag{5.4.10}$$

是 (5.4.7) 的解, 因而 $P(t)$ 和 L 对易. 结果, $P(t)$ 也是

$$\frac{\mathrm{d}P}{\mathrm{d}t}=LP,\quad P(0)=I \tag{5.4.11}$$

的解. 称 (5.4.7) 和 (5.5.11) 为 Markov 链的**主方程**.

例 5.6 设 $\alpha\in(0,+\infty)$, 并考虑连续时间 Markov 链, 其生成子为

$$L=\begin{pmatrix}-\alpha & \alpha\\ \alpha & -\alpha\end{pmatrix}.$$

计算 L 的特征值和特征函数, 将其分解为

$$L=U\begin{pmatrix}-2\alpha & 0\\ 0 & 0\end{pmatrix}U^{\mathrm{T}},$$

其中

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}.$$

矩阵 $L=U\Lambda U^{\mathrm{T}}$($\Lambda$ 为对角矩阵, U 为正交矩阵) 的指数矩阵定义如下:

$$\mathrm{e}^{tL}=U\mathrm{e}^{t\Lambda}U^{\mathrm{T}}.$$

用这个定义, 可以计算 L 生成的半群:

$$\mathrm{e}^{tL}=\frac{1}{2}\begin{pmatrix}1+\mathrm{e}^{-2\alpha t} & 1-\mathrm{e}^{-2\alpha t}\\ 1-\mathrm{e}^{-2\alpha t} & 1+\mathrm{e}^{-2\alpha t}\end{pmatrix}.$$ □

设 $\rho(t)=(\rho_0(t),\rho_1(t),\cdots)^{\mathrm{T}}$ 是矩阵 $P(t)$ 第 i 行的转置, 即 $\rho_j(t)=p_{ij}(t)$ 列向量是概率, 该概率是一个在时间 t 上起始于状态 i, 结束于状态 $j\,(j\in\mathcal{I})$ 的系统. 设 e_i 为第 i 个单位向量, 除第 i 个元素是 1, 其他均为 0, 直接由 (5.4.7) 得出下面的结论.

定理 5.7 概率向量 ρ 满足

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=L^{\mathrm{T}}\rho,\quad \rho(0)=e_i. \tag{5.4.12}$$

如果 Markov 链的初始状态是随机的, 并且在 Markov 链中独立选取转移概率向量 $\rho(0)=\rho_0$, 则

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=L^{\mathrm{T}}\rho,\quad \rho(0)=\rho_0. \tag{5.4.13}$$

证明　第一个结论可由 (5.4.7) 得出, 设 $\rho^{(i)}$ 为 (5.4.12) 的解. 如果初值条件是随机的, 其中 $\rho(0)=\rho_0$, 向量 ρ_0 第 i 个分量为 $\rho_{0,i}$, 则

$$\rho(t)=\sum_i \rho_{0,i}\rho^{(i)}(t).$$

对上述等式两边求导, 并由 (5.4.12) 得到 (5.4.13). □

方程 (5.4.13) 类似于 Liouville 方程和 Fokker-Plank 方程的 Markov 链, 其中 Liouville 方程和 Fokker-Plank 方程分别在第 4 章和第 6 章中进行了论述, 称之为**向前方程**.

设 $\phi:\mathcal{I}\mapsto\mathbb{R}$ 是定义在状态空间上的实值函数, 其可用元素为 $\phi_j(j\in\mathcal{I})$ 的向量表示. 定义 $v(t)=(v_0(t),v_1(t),\cdots)^{\mathrm{T}}$, 其第 i 个元素为

$$v_i(t)=\mathbb{E}\left(\phi_{z(t)}\middle|\, z(0)=i\right),$$

其中 $\mathbb{E}$ 定义为关于 Markov 转移概率的期望.

定理 5.8　期望的向量 v 满足方程

$$\frac{\mathrm{d}v}{\mathrm{d}t}=Lv,\quad v(0)=\phi. \tag{5.4.14}$$

证明　依据转移密度, 函数 $v_i(t)$ 可简记为

$$v_i(t)=\sum_j p_{ij}(t)\,\phi_j. \tag{5.4.15}$$

如果设 $\phi=(\phi_0,\phi_1,\cdots)^{\mathrm{T}}$, 则式 (5.4.15) 可写为向量形式 $v(t)=P(t)\phi$. 对时间求导并使用主方程 (5.4.11) 得到结果.

方程 (5.4.14) 类似于特征线法和向后 Kolmogorov 方程的 Markov 链, 特征线法和向后 Kolmogorov 方程分别在第 4 章和第 6 章中进行了论述, 称之为**向后方程**.

5.5　存在唯一性

连续时间 Markov 链有样本路径, 在 $\mathcal{I}$ 上该路径从一个状态跳跃到另一个. 在条件 (5.4.2) 下 (当条件不成立时, 跳跃速率依赖于当前状态的构造是有可能出现的) 构造跳跃链可以说明这些. 记住这一点, 很自然得到下面的定义.

定义 5.9　如果跳跃时间在有限时间上不累积, 则连续时间 Markov 链是**不溢出的**, 理解确保不溢出的条件是重要的. 证明条件 (5.4.2) 是充分的, 由离散时间的 Markov 链得到一个序列 $\{z_n\}_{n\geqslant 0}$, 其中 (5.4.3) 给出该 Markov 链的转移矩阵, 并且 $\lambda>l^*$. 然后由 (5.3.1) 得到与转移矩阵 P 的选择有关的跳跃链, 其中 $t_{n+1}=t_n+\tau_n$ 且 τ_n 是独立同分布随机变量, 其分布为 $\exp(\lambda)$.

定理 5.10 如果 (5.4.2) 成立, 则 Markov 链是不溢出的.

证明 确定 Markov 链相应的生成子 L, 此生成子具有跳跃链. 设

$$\varsigma = \sum_{n=0}^{\infty} \tau_n$$

和 $T_n = \lambda \tau_n$, 并且注意到 $\{T_n\}$ 构成一个独立同分布序列, 其中 $T_1 \sim \exp(1)$. 由强大数定理 (例 3.26),

$$\frac{1}{N+1} \sum_{n=0}^{N} T_n \to 1 \text{ a.s.},$$

因而

$$\varsigma := \lim_{N \to \infty} \frac{1}{\lambda} \sum_{n=0}^{N} T_n = \infty \text{ a.s..}$$

得到结论.

值得注意的是只要 $\mathcal{I}$ 是有限的, 条件 (5.4.2) 就均满足.

5.6 遍 历 性

对于常微分方程, Markov 链的遍历性与不变测度的存在唯一性有关. 把时间平均收敛到一个值的性质 (该值独立于初值条件并由唯一不变测度决定)、生成子零空间的特性以及生成子的伴随算子这三方面的与遍历性定义联系起来. 在可数状态空间 Markov 链中, 不变测度具有矢量性, 后面将定义不变分布.

简言之, 假定 $\mathcal{I}$ 是有限集合, 从讨论离散时间 Markov 链来开始论述. 通过 (5.2.1), 矩阵 $(P - I)$ 有非空的零空间, 此零空间包含有常数向量, 则其转置也有非空的零空间. 结果存在向量 ρ^∞, 使得

$$P^{\mathrm{T}} \rho^\infty = \rho^\infty. \tag{5.6.1}$$

实际上有下面的定理.

定理 5.11 如果矩阵 P 的所有特征值都在闭合单位圆上, 则选择向量 ρ^∞, 使得其所有元素都是非负的且 $\langle \rho^\infty, 1 \rangle = 1$.

向量 ρ^∞ 就是**不变分布**, 注意其在 $\mathcal{I}$ 上定义了一个不变测度. 在常微分方程中, 遍历性与使这个不变分布是唯一的有关, 通过下面的条件可直接确保这一点.

定义 5.12 除了在 1 上有单特征值, 如果矩阵 P 的谱严格存在于单位圆内部, 相当于一个严格不变正态分布, 则称离散时间 Markov 链具有遍历性.

现在在 $\mathcal{I}$ 上考虑连续时间 Markov 链. 由生成子 L 的性质得

$$L1 = 0, \quad L^{\mathrm{T}} \rho^\infty = 0. \tag{5.6.2}$$

实际上, 使用 (5.4.3) 得到下面的定理.

定理 5.13　如果 L 的所有特征值都在左半平面, 则选择向量 ρ^∞, 使得其所有元素都是非负且 $\langle \rho^\infty, 1\rangle = 1$.

同样, 称向量 ρ^∞ 为**不变分布**. 在离散时间 Markov 链里, 其在 $\mathcal{I}$ 上定义了一个不变测度, 并且遍历性与这个不变分布是唯一的有关.

定义 5.14　除了在 0 上有单特征值, 如果矩阵 L 的谱严格存在于左半平面内, 相当于一个严格不变正态分布, 则称离散时间 Markov 链具有遍历性.

下面的定理描述了遍历连续时间 Markov 链的特性, 在本书剩余部分会用到该定理.

定理 5.15　在有限状态空间 $\mathcal{I}$ 上, 遍历连续时间 Markov 链满足下面 5 个性质:

(i) $\mathcal{N}(L) = \operatorname{span}\{1\}$;

(ii) $\mathcal{N}\left(L^{\mathrm{T}}\right) = \operatorname{span}\{\rho^\infty\}, \rho^\infty(i) > 0, \forall i \in \mathcal{I}$;

(iii) $\exists C, \lambda > 0$, 使得向前方程 (5.4.13) 的解满足

$$|\rho(t) - \rho^\infty|_1 \leqslant C\mathrm{e}^{-\lambda t}, \quad \forall t > 0;$$

(iv) $\exists C, \lambda > 0$, 使得向后方程 (5.4.14) 的解满足

$$|v(t) - \langle \rho^\infty, \phi\rangle 1|_\infty \leqslant C\mathrm{e}^{-\lambda t}, \quad \forall t > 0;$$

(v)

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi_{z(t)}\mathrm{d}t = \langle \rho^\infty, \phi\rangle \quad \text{a.s..}$$

由于状态空间是有限维的, 则用任意范数收敛结果都是成立的. 然而, 从概率论的观点来看, 前面讲的选择方法是顺理成章的.

注意: 通过选取 $\phi = e_i$, 从最终的结果得出 ρ^∞ 的第 i 个分量可被认为是时间比例, 该时间比例表示在 $t \in [0, \infty)$ 上 Markov 链的任意轨迹在状态 i 上掠过的时间, 这类似于前面章节中常微分方程例子, 尤其是式 (4.4.2). 还注意到在常微分方程中, 遍历性还与时间平均有关, 该时间平均收敛于用不变分布得到的平均, 因而遍历性独立于初始条件.

例 5.16　继续例 5.6. 注意 L 是由 $(1,1)^{\mathrm{T}}$ 所张成的一维零空间, 同样, L^{T} 也具有该性质, 则 $\rho^\infty = (1/2, 1/2)^{\mathrm{T}}$. 显然,

$$\lim_{t\to\infty} \mathrm{e}^{tL} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

而且极限依指数快速收敛. 由于 $\rho(0)$ 是一个分布 (见例 3.1), 得到 $\rho(t) = \mathrm{e}^{tL^T}, \rho(0)$ 依指数快速收敛于 ρ^∞. 因为 L 是对称的, $v(t)$ 同样依指数收敛. □

5.7 讨论和参考

优秀的 Markov 链入门教科书见文献 [235]; 更为深入的论著见文献 [303]; 教科书 [94] 里可以找到大量关于 Markov 过程, 包括 Markov 链的一般的材料; 论文 [122] 通过跳跃链的样本路径描述了模拟连续时间 Markov 链的简单算法; 书 [299] 从计算线性代数角度描述了 Markov 链; 书 [227] 对 Markov 链遍历性进行一个非常全面的论述.

由非负矩阵 (矩阵元素均为非负) 的一般理论得出在离散时间情况下, 非负不变向量的存在性. 用矩阵 P^{T} 简化特征对 $(\rho^{\infty},1)$, 并且对所有 i 确保 $\rho^{\infty}(i)>0$ 等价于要求矩阵是不可约的: 对于任意偶 $i,j\in\mathcal{I}$, 存在一个 $n\in\mathbb{Z}^{+}$, 使得 $(P^{n})_{ij}>0$. 要使得链具有遍历性就要求其具有非周期性和不可约性. 非周期性和不可约性暗含了存在 $n\in\mathbb{Z}^{+}$, 使得 P^{n} 的所有元素严格为正, 这意味着存在一个 $n\in\mathbb{Z}^{+}$ 有

$$\mathbb{P}(z_{n}=j|z_{0}=i)>0,\quad \forall(i,j)\in\mathcal{I}\times\mathcal{I},$$

因而存在正整数 n, 使得从 $\mathcal{I}$ 的任意起点, 利用正概率经过 n 步计算能够到达 $\mathcal{I}$ 中的任意项. 动态轨迹应当可以到达状态空间的全部区域, 这是另外一个思考遍历性的方法, 该方法很有用且具有探索性①. 非负矩阵的这些结论及其谱在 [115] 中可以找到; 注意关于遍历性不可约的例子以及不可约性和非周期性混合的例子见教科书 [327].

使用跳跃链, 在有限状态空间里, 直接可以将这些遍历性的结论延伸到具有连续时间的情形上. 回想对于每个在有限状态空间 $\mathcal{I}$ 上生成子为 L 的连续时间 Markov 链, 可以将一个跳跃链与 (5.4.3) 给出的转换矩阵联系起来, 其中任意 $\lambda>l^{*}$, l^{*} 由 (5.4.2) 得到. 由离散时间过程 (转换矩阵为 P) 非负不变矢量的存在性可以得到连续时间过程非负不变矢量的存在性. 这是因为 P 和 L 有相同的特征向量, 并且特征值 $\eta(P)$ 和 $\eta(L)$ 有下列关系:

$$\eta(P)=1+\lambda^{-1}\eta(L).\tag{5.7.1}$$

特别地, 对于不变向量, 如果 $\eta(P)=1$, 则 $\eta(L)=0$. 对于连续时间过程遍历性的判别依据简化为下列条件: (5.4.3) 给出的矩阵 P 是不可约的, 其中 $\lambda>l^{*}$, l^{*} 由 (5.4.2) 得到. 那么矩阵 P 的谱存在于闭合单位圆内部, 在 1 上有单特征值. 由 (5.7.1) 得到的矩阵 L 严格在左半平面上有谱, 并且在 0 上有单特征值, 则在连续时间情形中, 就不要求有非周期性这个概念了.

① 实际上, 这就是单词 ergodic 的词根, 即 ergodic $=\varepsilon\rho\gamma o+o\delta o\varsigma$. 这是 19 世纪 70 年代 Boltzmann 在其关于气体分子运动论的著作者引入的概念.

5.8 练 习

1. 证明跳跃链 (5.3.1) 满足 Markov 特性.

2. 给出 $\mathbb{Z}^d$ 上对称随机行走的定义, 其中 d 是任意的.

3. 给出 $\mathbb{Z}$ 上反对称随机行走的定义.

4. 考虑具有下面运输矩阵的离散时间 Markov 链:

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix},$$

其中 $\alpha, \beta \in [0,1]$. 找出不变密度 ρ^∞. 在什么条件下, Markov 链具有遍历性?

5. 假设

$$p_{ij}\pi_i = p_{ji}\pi_j, \quad \forall i, j \in \mathcal{I},$$

则称随机矩阵 P 和分布 $\boldsymbol{\pi}$**平衡**. 证明对于 P, π 是一个不变分布. (加强该条件与使 Markov 链是可逆的有关.)

6. 考虑连续时间 Markov 链, 其生成子为

$$L = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}.$$

找到不变密度 ρ^∞. 在什么条件下, Markov 链具有遍历性?

7. 由前面的例子, 使用标准跳跃链, 用一个数值解法来模拟连续时间 Markov 链.

第 6 章　随机微分方程

6.1　格　　局

Gaussian 白噪声是指均值为零、具有相关性矩阵 $\delta(t-s)I$ 的 Gaussian 过程. 实际上, 这是对相关时间很短的随机过程的一种理想化. 本章中, 给出一些关于随机微分方程 (由白噪声推导得到) 的背景资料, 该内容在书的剩余部分中要用到. 定义 $W(t)$ 为一个标准 m 维 Brownian 运动, $h:\mathcal{Z}\to\mathbb{R}^d$ 是一个光滑的向量值函数, $\gamma:\mathcal{Z}\to\mathbb{R}^{d\times m}$ 是一个光滑矩阵值函数. 接下来, 特别地, 取 $\mathcal{Z}=\mathbb{T}^d,\mathbb{R}^d$ 或 $\mathbb{R}^l\oplus\mathbb{T}^{d-l}$, 考虑 Itô 随机微分方程

$$\frac{\mathrm{d}z}{\mathrm{d}t}=h(z)+\gamma(z)\frac{\mathrm{d}W}{\mathrm{d}t},\quad z(0)=z_0. \tag{6.1.1}$$

将 $\mathrm{d}W/\mathrm{d}t$ 看成是 Gaussian 白噪声. 这种过程只有作为一个分布时才成立. 因此, 准确理解 (6.1.1) 就是对所有 $z(t)\in C(\mathbb{R}^+,\mathcal{Z})$, 理解如下积分方程:

$$z(t)=z_0+\int_0^t h(z(s))\,\mathrm{d}s+\int_0^t\gamma(z(s))\,\mathrm{d}W(s). \tag{6.1.2}$$

为了理解这个方程, 需要定义 $\mathrm{d}W(s)$ 的随机积分, 使用在第 3 章中定义的随机积分的 Itô 解释. 因为这样做在概念上是方便的, 所以通常将 (6.1.1) 的未积分形式记作随机微分方程, 任何时候都应将其解释为是 (6.1.2) 的简略形式. 有时将 (6.1.1) 中的函数 h 看成是一种漂移, γ 看成是扩散系数.

在后面, 通常认为 $z=\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}$, 其中 $x\in\mathcal{X},y\in\mathcal{Y}$. 如果 $\mathcal{Z}=\mathbb{T}^d\left(\mathbb{R}^d\right)$, 则 $\mathcal{X}=\mathbb{T}^l(\mathbb{R}^l)$, $\mathcal{Y}=\mathbb{T}^{d-l}(\mathbb{R}^{d-l})$. 如果 $\mathcal{Z}=\mathbb{R}^l\oplus\mathbb{T}^{d-l}$, 则 $\mathcal{X}=\mathbb{R}^l$, $\mathcal{Y}=\mathbb{T}^{d-l}$. 当在环面上考虑方程 (4.1.1) 时, 就像第 4 章中那样, 这仅仅是下列情形的一种简写形式: h 和 γ 在 z(或 y) 上具有周期性, 并且通过令环面各方向上模为 1, 将 z(或 y) 看成是环面的元素. 有时, 在环面上考虑微分算子, 这时算子自动具有周期边界条件.

6.2 节 ~6.4 节分别考虑解的存在唯一性、生成子、Kolmogorov 方程、Fokker-Planck 方程以及随机微分方程的遍历性. 在 6.5 节中, 对基本理论进行延伸并给出相关参考文献.

6.2　存在唯一性

通过 (6.1.1) 的解, 用下面的性质来说明一个 $\mathcal{Z}$ 值随机过程 $\{z(t)\}$, 其中 $t\in$

$[0, T]$:

(i) z 是连续的且适应 $\mathcal{F}_t$ 鞅, 其中 Brownian 运动 $W(t)$ 生成其代数;

(ii) $h(z(\cdot)) \in L^1((0,T)), \gamma(z(\cdot)) \in L^2((0,T))$;

(iii) 对每个 $t(t \in [0,T])$, 方程 (6.1.1) 依概率为 1 成立.

如果任意两个解 $x_i(t)\,(i=1,2)$ 满足

$$\mathbb{P}(x_1(t) = x_2(t), \forall t \in [0,T]) = 1,$$

则称解是唯一的.

在定理 4.2 中, 已经证明了对所有 Lipschitz 向量场 h, 常微分方程解的存在唯一性 (即在 (6.1.1) 中, 当 $\gamma \equiv 0$ 时). 当 $\gamma \neq 0$ 时, 存在一个类似的定理. 对于常微分方程, 当找到一个解的先验性边界时, 条件就可以减弱, 但是为了说明问题, 简单地建立如下定理:

定理 6.1　假定 $h(\cdot)$ 和 $\gamma(\cdot)$ 在 $\mathcal{Z}$ 上是完全 Lipschitz 的, 并且 z_0 是一个独立于 Brownian 运动 $W(t)$ 的随机变量, 满足

$$\mathbb{E}|z_0|^2 < \infty,$$

则随机微分方程 (6.1.1) 有唯一解 $z \in C(\mathbb{R}^+; \mathcal{Z})$, 其中

$$\mathbb{E}\left[\int_0^T |z(t)|^2 \mathrm{d}t\right] < \infty, \quad \forall T < \infty,$$

而且随机微分方程的解是一个 Markov 过程.

用两个备注归纳本节, 这两个备注在后面章节有很重要的作用.

注 6.2　(6.1.1) 的 Stratonovich 模拟为

$$\frac{\mathrm{d}z}{\mathrm{d}t} = h(z) + \gamma(z)_0 \frac{\mathrm{d}W}{\mathrm{d}t}, \quad z(0) = z_0. \tag{6.2.1}$$

通过这个方程说明 $z \in C(\mathbb{R}^+, \mathcal{Z})$ 满足积分方程

$$z(t) = z(0) + \int_0^t h(z(s))\,\mathrm{d}s + \int_0^t \gamma(z(s))_0 \mathrm{d}W(s). \tag{6.2.2}$$

使用式 (3.4.6) 和 (3.4.9), 如果假定 $\gamma(z)$ 是可微的, 显然满足 Stratonovich 随机微分方程的 z 也满足 Itô 随机微分方程

$$\frac{\mathrm{d}z}{\mathrm{d}t} = h(z) + \frac{1}{2}\nabla\cdot\left(\gamma(z)\gamma(z)^{\mathrm{T}}\right) - \frac{1}{2}\gamma(z)\nabla\cdot\left(\gamma(z)^{\mathrm{T}}\right) + \gamma(z)\frac{\mathrm{d}W}{\mathrm{d}t}, \quad z(0) = z_0, \tag{6.2.3}$$

见练习第 1 题.

在大部分应用中，白噪声是对相关时间很短的随机过程的一种理想化. 此时，随机微分方程的 Stratonovich 解释尤为重要，因为有时这个解释会以极限形式出现，该极限可通过对白噪声进行光滑逼近得到①. 另一方面，作为一个数学对象，由 Itô 积分得到的鞅原理使得这个 Stratonovich 解释更加重要，因而两种观点之间的相互转化是非常有用的. □

注 6.3 Brownian 运动定义 3.5 暗含如下有趣的标度性质：

$$\{W(ct)\,|t \geqslant 0\} = \left\{\sqrt{c}W(t)\,|t \geqslant 0\right\},$$

其中前面的内容应当解释为在 $C\left([0,\infty),\mathbb{R}^d\right)$ 依律成立. 由此得出，如果 $s = ct$，则

$$\frac{\mathrm{d}W}{\mathrm{d}s} = \frac{1}{\sqrt{c}}\frac{\mathrm{d}W}{\mathrm{d}t}.$$

因此，如果在 (6.1.1) 中变时间尺度 $s = ct$，则得到方程

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \frac{1}{c}h(z) + \frac{1}{\sqrt{c}}\gamma(z)\frac{\mathrm{d}W}{\mathrm{d}s}, \quad z(0) = z_0.$$

(一般来说，积分方程是对其进行精确的理解.) 注意到因为由扩散系数得到的二次项 (6.3.1)，特别情况，当取 $s = ct$ 时，下节中定义的随机微分方程变换 (Fokker-Planck 方程) 是一种标准变换. □

6.3 生 成 子

在随机微分方程 (6.1.1) 中给出函数 $\gamma(z)$，定义

$$\Gamma(z) = \gamma(z)\gamma(z)^{\mathrm{T}}, \tag{6.3.1}$$

则生成子$\mathcal{L}$ 定义为

$$\mathcal{L}v = h\cdot\nabla v + \frac{1}{2}\Gamma : \nabla\nabla v. \tag{6.3.2}$$

按照第 3 章的定义，具有合适定义域的算子是 (6.1.1) 给出的 Markov 过程的生成子，则正规 L^2 的伴随算子 $\mathcal{L}^*$ 为

$$\mathcal{L}^*v = -\nabla\cdot(hv) + \frac{1}{2}\nabla\cdot\nabla\cdot(\Gamma v).$$

例 6.4 (i) 在 $\mathbb{R}^d$ 上考虑随机微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma\frac{\mathrm{d}W}{\mathrm{d}t},$$

① 在一些不是很严格的常规假设下，这个结论通常在一维情形下是成立的，而在更高维的情形下就不成立了：除了 Stratonovich 随机积分，当去除正则化时，在极限中会出现一个关于扩散矩阵 $\gamma(z)$ 的列向量之间的通信子的附加漂移，见 11.7.7 节中的例子以及 11.8 节的讨论.

其解为 $\mathbb{R}^d$ 上的一个协方差为 $\sigma^2 I$ 的 Brownian 运动.

Markov 过程 $x(t)$ 的生成子为

$$\mathcal{L}\phi = \frac{\sigma^2}{2}\Delta\phi,$$

其形式在 $L^2(\mathbb{R}^d)$ 上是自共轭的,

$$\mathcal{L}^*\phi = \frac{\sigma^2}{2}\Delta\phi.$$

(ii) 在 $\mathbb{R}^d$ 上考虑 Ornsterin-Uhlenbeck(OU) 随机微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\alpha x + \sqrt{2\lambda}\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中 α, λ 为正常数. OU 过程的生成子是

$$\mathcal{L}\phi = -\alpha x \cdot \nabla\phi + \lambda\Delta\phi,$$

其 $L^2(\mathbb{R}^d)$ 伴随算子为

$$\mathcal{L}^*\phi = \nabla \cdot (\alpha x\phi) + \lambda\Delta\phi.$$

事实上, 通过在加权 L^2 空间上对关于 OU 过程不变测度的密度进行积分 (见练习第 4 题) 可知 $\mathcal{L}$ 是自共轭的.

(iii) 在 $\mathbb{R}$ 上的几何 Brownian 运动定义为如下一维随机微分方程的解:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu x + \sigma x\frac{\mathrm{d}W}{\mathrm{d}t},$$

此处 $\mu \in \mathbb{R}$ 且 $\sigma > 0$. 几何 Brownian 运动的生成子为

$$\mathcal{L}\phi = \mu x\frac{\mathrm{d}\phi}{\mathrm{d}x} + \frac{\sigma^2 x^2}{2}\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2},$$

其 L^2 伴随算子是

$$\mathcal{L}^*\phi = -\frac{\mathrm{d}}{\mathrm{d}x}(\mu x\phi) + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\sigma^2 x^2}{2}\phi\right).$$

□

下面的 Itô**公式**是关于函数 $V: \mathcal{Z} \to \mathbb{R}$ 在时间上变化率的基本结果, 该函数由求解一个 $\mathcal{Z}$ 值随机积分方程得到. 具有启发式地得到下面的结果:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V(z(t))) = \mathcal{L}V(z(t)) + \left\langle \nabla V(z(t)), \gamma(z(t))\frac{\mathrm{d}W}{\mathrm{d}t}\right\rangle.$$

这类似于 (4.3.2) 中常微分方程的情况. 注意到如果 W 是一个光滑时变函数, 则这个公式就可能不正确: 由于 Brownian 运动缺乏光滑性, 导致在 $\mathcal{L}V$ 上存在一个与 Γ 成比例的附加项.

至于随机微分方程 (6.1.1) 本身, 可用积分形式对 V 变化率的表达式进行精确理解.

引理 6.5 (Itô 公式) 假设定理 6.1 的条件成立, 设 $x(t)$ 是 (6.1.1) 的解且 $V \in C^2(\mathcal{Z}, \mathbb{R})$, 则过程 $V(z(t))$ 满足

$$V(z(t)) = V(z(0)) + \int_0^t \mathcal{L}V(z(s))\,\mathrm{d}s + \int_0^t \langle \nabla V(z(s)), \gamma(z(s))\,\mathrm{d}W(s)\rangle.$$

在常微分方程的条件下, 公式很容易推广到向量值函数 $V: \mathcal{Z} \to \mathbb{R}^n$ 上.

设 $\phi: \mathcal{Z} \mapsto \mathbb{R}$, 考虑方程

$$v(z,t) = \mathbb{E}(\phi(z(t)) | z(0) = z), \tag{6.3.3}$$

其中期望是对于所有 Brownian 的驱使路径. 通过在 Itô 公式中求期望, 消去随机积分, 下面可推出引理 6.5 的重要结论.

定理 6.6 假设 ϕ 足够光滑使得向后 **Kolmogorov 方程**

$$\begin{aligned} \frac{\partial v}{\partial t} &= \mathcal{L}v, \quad (z,t) \in \mathcal{Z} \times (0,\infty), \\ v &= \phi, \quad (z,t) \in \mathcal{Z} \times \{0\} \end{aligned} \tag{6.3.4}$$

存在唯一古典有界解[①] $v(x,t) \in C^{2,1}(\mathcal{Z} \times (0,\infty), \mathbb{R}) \cap C(\mathcal{Z} \times \mathbb{R}^+, \mathbb{R})$, 其中 $\mathcal{L}$ 由 (6.3.2) 给出. 而 v 由 (6.3.3) 得到, 其中 $z(t)$ 是 (6.1.2) 的解.

这类似于 Markov 链的向后方程 (5.4.14). 如果 (6.1.1) 中 $\gamma \equiv 0$, 使得动力系统是确定性的, 并且由 φ^t 是 $\mathcal{Z}$ 上的流, 得到 $z(t) = \varphi^t(z(0))$, 则 Kolmogorov 方程 (6.3.4) 简化为双曲型方程 (4.3.5), 其特征曲线是常微分方程 (4.1.1) 的积分曲线.

例 6.7 考虑例 6.4 中的三种随机微分方程, 相应的向后 Kolmogorov 方程为

Brownian 运动:

$$\frac{\partial v}{\partial t} = \frac{\sigma^2}{2}\Delta v;$$

OU 过程:

$$\frac{\partial v}{\partial t} = -\alpha x \cdot \nabla v + \lambda \Delta v;$$

几何 Brownian 运动:

$$\frac{\partial v}{\partial t} = \mu x \frac{\partial v}{\partial x} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 v}{\partial x^2}.$$

□

由定理 6.6 可以直接得到一个密度传播方程. 在接下来的定理中推导此方程时, 会用到时变偏微分方程解的半群这一概念, 见 7.5 节.

① 见第 7 章.

定理 6.8 考虑方程 (6.1.2), 其中 z_0 是独立于 w, 密度为 $\rho_0(z)$ 的随机变量. 假定 $z(t)$ 的分布律有密度 $\rho(z,t) \in C^{2,1}(\mathcal{Z}\times(0,\infty),\mathbb{R})\bigcap C(\mathcal{Z}\times\mathbb{R}^+,\mathbb{R})$, 则 ρ 满足**Fokker-Planck 方程**①

$$\frac{\partial\rho}{\partial t} = \mathcal{L}^*\rho, \quad (z,t)\in\mathcal{Z}\times(0,\infty), \tag{6.3.5a}$$

$$\rho = \rho_0, \quad z\in\mathcal{Z}\times\{0\}. \tag{6.3.5b}$$

证明 $\mathbb{E}^\mu$ 定义为关于乘积测度的平均, 该测度由 $z(0)$ 上密度为 ρ_0 的测度 μ 以及随机微分方程本身独立驱动 Wiener 测度得到. 由前面的结论, 在分布密度为 $\rho_0(z_0)$ 的随机 z_0 上求平均值可以发现

$$\begin{aligned}\mathbb{E}^\mu(\phi(z(t))) &= \int_{\mathcal{Z}} v(z_0,t)\rho_0(z_0)\,\mathrm{d}z_0 = \int_{\mathcal{Z}} v(z,t)\rho_0(z)\,\mathrm{d}z\\ &= \int_{\mathcal{Z}}\left(\mathrm{e}^{\mathcal{L}t}\phi\right)(z)\,\rho_0(z)\,\mathrm{d}z = \int_{\mathcal{Z}}\left(\mathrm{e}^{\mathcal{L}^*t}\rho_0\right)(z)\,\phi(z)\,\mathrm{d}z.\end{aligned}$$

对于所有足够光滑的 ϕ, 定理 6.6 成立. 但是如果 $z(t)$ 的密度为 $\rho(z,t)$, 也可得

$$\mathbb{E}^\mu(\phi(z(t))) = \int_{\mathcal{Z}}\rho(z,t)\,\phi(z)\,\mathrm{d}z.$$

令这两个表达式在时间 t 上的期望相等得

$$\int_{\mathcal{Z}}\left(\mathrm{e}^{\mathcal{L}^*t}\rho_0\right)(z)\,\phi(z)\,\mathrm{d}z = \int_{\mathcal{Z}}\rho(z,t)\,\phi(z)\,\mathrm{d}z.$$

用密度理论将密度延伸到所有 $\phi\in L^2(\mathcal{Z})$ 上, 从而由前面的方程可得

$$\rho(z,t) = \left(\mathrm{e}^{\mathcal{L}^*t}\rho_0\right)(z).$$

对方程求导得 (6.3.5a). 设 $t=0$ 为初值条件 (6.3.5b), 从而 ρ 是初值问题 (6.3.5) 的解. □

Fokker-Planck 方程依次类似于 Markov 链的向前方程 (5.4.12) 和常微分方程的 Liouville 方程.

注 6.9 事实上 Fokker-Planck 方程的解是概率测度的密度意味着

$$\int_{\mathcal{Z}}\rho(z,t)\,\mathrm{d}z = 1, \quad \rho(z,t)\geqslant 0, \quad \forall t\in\mathbb{R}^+.$$

因为 $\mathcal{L}^*$ 是散度型算子, 由 (6.3.5) 得出积分守恒律, 而且概率表示暗含如果 ρ 在 $t=0$ 时是非负的, 则其对所有正时间都是非负的. □

① 有时称其为向前 **Kolmogorov** 方程.

注 6.10 在随机微分方程 (6.1.1) 的初始条件是确定的情况下, 即 $z=z_0$, Fokker-Planck 方程的初始条件就变成一个 delta 函数 (Dirac 质量)

$$\rho(z,t=0)=\delta(z-z_0).$$

Fokker-Planck 方程的解可以拓展到适用于那些非正则性情况. □

例 6.11 考虑例 6.4 中的三种随机微分方程, 相应的 Fokker-Planck 方程为

Brownian 运动:

$$\frac{\partial\rho}{\partial t}=\frac{\sigma^2}{2}\Delta\rho;$$

OU 过程:

$$\frac{\partial\rho}{\partial t}=\nabla\cdot(\alpha x\rho)+\lambda\Delta\rho;\tag{6.3.6}$$

几何 Brownian 运动:

$$\frac{\partial\rho}{\partial t}=-\frac{\partial}{\partial x}(\mu x\rho)+\frac{\partial^2}{\partial x^2}\left(\frac{\sigma^2x^2}{2}\rho\right).$$ □

注意: 常数在生成子 $\mathcal{L}$(由 (6.3.2) 得到) 的零空间上, 其暗含空间维数 (零空间 $\mathcal{L}$) $\geqslant 1$. 假设生成子 $\mathcal{L}$ 满足 Fredholm 变换 (见 7.2.3 小节), 也暗含了空间维数 (零空间 $\mathcal{L}^*$) $\geqslant 1$. 因此, 希望 $\mathcal{L}^*$ 有非平凡零空间.

设 ρ 为零空间 $\mathcal{L}^*$ 上的一个函数, 其为正且在 $\mathcal{Z}$ 上积分为 1, 则 ρ 一定是 Fokker-Planck 方程的稳态解,

$$\mathcal{L}^*\rho=0.\tag{6.3.7}$$

如果给定方程 (6.3.7) 适当的边界条件, 从而得到具有一个平稳分布的 ρ, 其是椭圆型偏微分方程 (6.3.7) 的解. 如果给定与算子 $\mathcal{L}$ 有关的**Dirichlet 形式**, 则有下面的重要结论. 为了简便起见, 在周期条件下把它提出来并证明它.

定理 6.12 在 $\mathbb{T}^d$ 上, 设 $\rho\in C^2_{\text{per}}(\mathbb{T}^d)$ 是任意 Fokker-Planck 方程的稳态解, 并且具有周期边界条件. 设 $f\in C^2_{\text{per}}(\mathbb{T}^d)$, 则

$$\int_{\mathbb{T}^d}(-\mathcal{L}f(2))f(z)\rho(z)\mathrm{d}z=\frac{1}{2}\int_{\mathbb{T}^d}(\nabla f(z)\cdot\varGamma(z)\nabla f(z))\rho(z)\mathrm{d}z\tag{6.3.8}$$

$$=\frac{1}{2}\int_{\mathbb{T}^d}|\gamma(z)^{\mathrm{T}}\nabla f(z)|^2\rho(2)\mathrm{d}z.\tag{6.3.9}$$

证明 有

$$\begin{aligned}\mathcal{L}^*(f\rho)=&-\nabla\cdot(h\rho f)+\frac{1}{2}\nabla\cdot\nabla\cdot(\varGamma\rho f)\\=&(\mathcal{L}^*\rho)f+(-\mathcal{L}f)\rho+\nabla\rho\cdot\varGamma\nabla f+\rho\varGamma:\nabla\nabla f+\nabla f\cdot(\nabla\cdot\varGamma)\rho\\=&(-\mathcal{L}f)\rho+\nabla\rho\cdot\varGamma\nabla f+\rho\varGamma:\nabla\nabla f+\nabla f\cdot(\nabla\cdot\varGamma)\rho.\end{aligned}$$

设 $g \in C_{\mathrm{per}}^{2}\left(\mathbb{T}^{d}\right)$, 事实上, 在 $L^{2}\left(\mathbb{T}^{d}\right)$ 中 $\mathcal{L}$ 和 $\mathcal{L}^{*}$ 互为伴随算子, 应用散度定理, 使用前面的恒等式进行如下运算:

$$\begin{aligned}\int_{\mathbb{T}^{d}}(\mathcal{L}g)f\rho\mathrm{d}z &= \int_{\mathbb{T}^{d}}g\mathcal{L}^{*}(f\rho)\,\mathrm{d}z \\ &= \int_{\mathbb{T}^{d}}\left(g(-\mathcal{L}f)\rho + g\nabla\rho\cdot\Gamma\nabla f\right)\mathrm{d}z \\ &\quad + \int_{\mathbb{T}^{d}}g\left(\nabla f\cdot(\nabla\cdot\Gamma)\rho + \rho\Gamma:\nabla\nabla f\right)\mathrm{d}z \\ &= \int_{\mathbb{T}^{d}}g(-\mathcal{L}f)\rho\mathrm{d}z - \int_{\mathbb{T}^{d}}(\nabla g\cdot\Gamma\nabla f)\rho\mathrm{d}z. \end{aligned}\tag{6.3.10}$$

方程 (6.3.10) 暗含

$$\int_{\mathbb{T}^{d}}(-\mathcal{L}g)f\rho\mathrm{d}z + \int_{\mathbb{T}^{d}}g(-\mathcal{L}f)\rho\mathrm{d}z = \int_{\mathbb{T}^{d}}(\nabla g\cdot\Gamma\nabla f)\rho\mathrm{d}z. \tag{6.3.11}$$

设 $g=f$, 由前面的方程得到方程 (6.3.8). □

大致来说, 前面关于 Dirichlet 形式的结论表明在一个适当的加权 L^{2} 空间上, $-\mathcal{L}$ 是一个正算子. 的确, 设 ρ 在 $\mathcal{Z}$ 上严格为正, 并且定义测度 $\mu(\mathrm{d}x)=\rho(x)\,\mathrm{d}x$, 则引入如下具有内积形式的加权 Hilbert 空间 $L^{2}(\mu)$:

$$(a,b)_{\rho} = \int_{\mathbb{T}^{d}}a(z)\cdot b(z)\rho(z)\,\mathrm{d}z, \quad \|a\|_{\rho}^{2} = (a,a)_{\rho}^{2}, \tag{6.3.12}$$

则前面的定理表示为

$$(-\mathcal{L}f,f)_{\rho} = \frac{1}{2}\left\|\gamma^{\mathrm{T}}\nabla f\right\|_{\rho}^{2}. \tag{6.3.13}$$

注 6.13　在适当的函数条件下, 前面有关 Dirichlet 形式的结论可用于定义域为 $\mathcal{Z}$ 的其他情况, 而不仅仅只适用于环面. 特别地, 函数空间应当确保 $\mathcal{L}^{*}$ 是 $\mathcal{L}$ 的伴随, 并且保证使用散度定理计算可以得到 (6.3.10). 比较典型地, 在非紧空间上通过在无限远处进行衰减假设, 这些条件都可以满足. □

例 6.14　例如 6.4(i), 在 $\mathbb{T}$ 上的 Brownian 运动有 $\rho(x)\equiv 1$, 使得 Lebesgue 测度是不变量. 此时, 通过分部积分 (6.3.13) 简化为积分形式

$$-\frac{\sigma^{2}}{2}\int_{0}^{1}\frac{\mathrm{d}^{2}f}{\mathrm{d}x^{2}}f\mathrm{d}x = \frac{\sigma^{2}}{2}\int_{0}^{1}\left|\frac{\mathrm{d}f}{\mathrm{d}x}\right|^{2}\mathrm{d}x.$$ □

6.4 遍　历　性

对于常微分方程和 Markov 链, 遍历性与唯一不变测度的存在性有关. 对于连续时间 Markov 链, 在随机微分方程中, 遍历性的特征是在伴随算子零空间上求一个唯一的 (通过标准化) 函数, 称之为不变分布或平稳分布.

非常有必要地, 在 $\mathcal{L}$ 的零空间上一定有常数

$$\mathcal{L}\mathbf{1} = 0.$$

通常记号 **1** 定义为一个不变且恒等于 1 的函数, 即在 L^p 上成立. 大体上说, 如果随机微分方程生成子的零空间中仅含有常数, 则称随机微分方程具有遍历性. 假设 $\mathcal{L}$ 满足 Fredholm 变换, 这等价于 Fokker-Planck 平稳方程存在一个唯一、标准归一的解

$$\mathcal{L}^*\rho^\infty = 0, \quad \inf_{z\in\mathcal{Z}} \rho^\infty > 0, \quad \int_{\mathcal{Z}} \rho^\infty(z)\,\mathrm{d}z = 1.$$

在这个例子中, 随机微分方程解函数的长时间平均等于相对**不变分布**$\rho^\infty(z)$ 的平均, 从而使得随机微分方程忽略其初始条件. 方程 (6.4.2) 准确地解释了这个思想, 在物理学中通常将其看成遍历性的定义. 除了与随机微分方程的单一样本路径有关外, 遍历性还与样本路径的集合有关, 这些样本路径是依赖于不同的噪声来实现的. 当 $t \to \infty$ 时, 对于遍历性随机微分方程, Fokker-Planck 方程 (6.3.5) 的解收敛于其不变分布, 见 (6.4.1).

给出一个随机微分方程, 想知道其是否具有遍历性. 当 $\mathcal{Z} = \mathbb{T}^d$ 时给出严格的结论. 在 $\mathbb{T}^d$ 上考虑随机微分方程 (6.1.1), 使得生成子 $\mathcal{L}$ 及其伴随 $\mathcal{L}^*$ 具有周期边界条件, 则

$$\mathcal{D}(\mathcal{L}) = \mathcal{D}(\mathcal{L}^*) = C^2_{\text{per}}\left(\mathbb{T}^d\right).$$

下面是在数学上对随机微分方程遍历性的准确表达, 它确保刚才描述的具有探索性的想法能够在周期性例子中实现. 下面用推广定义来表示连续时间 Markov 链遍历性的定义.

定义 6.15 如果除了在原点具有对应一个严格正特征函数 $\rho^\infty(z)$ 的简单特征值之外, 生成子的谱严格存在于左半平面, 则称随机微分方程 (6.1.1) 具有遍历性, 其中 $\mathcal{Z} = \mathbb{T}^d$.

接下来使用简化符号 $\rho(t)$ 和 $v(t)$ 分别定义 Fokker-Planck 的时变函数值解和向后 Kolmogorov 方程的解. 因此, 对每个固定的 t, 可将 $\rho(t)$ 和 $v(t)$ 视为在 Banach 空间上, 并且通过 L^p 范数测定它们的大小.

定理 6.16 在 $\mathbb{T}^d$ 上使 $\mathcal{L}$ 和 $\mathcal{L}^*$ 具有周期边界条件, 并且假定 $\Gamma(z)$ 在 $z = \mathbb{T}^d$ 上是一致严格正定的,

$$\exists\overline{\gamma} > 0: \ \langle\xi, \Gamma(z)\xi\rangle \geqslant \overline{\gamma}|\xi|^2, \forall\xi \in \mathbb{R}^d, z \in \mathbb{T}^d,$$

则随机微分方程 (6.1.1) 具有遍历性且满足下面 5 个特性:

- $\mathcal{N}(\mathcal{L}) = \text{span}\{\mathbf{1}\}$;

- $\mathcal{N}(\mathcal{L}^*) = \operatorname{span}\{\rho^\infty\},\ \inf_{z\in\mathbb{T}^d} \rho^\infty(z) > 0$;
- $\exists C, \lambda > 0$, 使得在任意 $z(0) \in \mathbb{T}^d$ 上具有初值 Dirac 质量的 Fokker-Planck 方程的解满足

$$\|\rho(t) - \rho^\infty\|_1 \leqslant C\mathrm{e}^{-\lambda t}, \quad \forall t > 0; \tag{6.4.1}$$

- $\exists C, \lambda > 0$, 使得初值为连续函数 ϕ 的向后 Kolmogorov 方程的解满足

$$\left\| v(t) - \left(\int_{\mathbb{T}^d} \phi(z)\rho^\infty(z)\,\mathrm{d}z\right) 1 \right\|_\infty \leqslant C\mathrm{e}^{-\lambda t}, \quad \forall t > 0;$$

- 对于所有 $\phi \in C(\mathbb{T}^d)$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(z(t))\,\mathrm{d}t = \overline{\phi} := \int_{\mathbb{T}^d} \phi(z)\rho^\infty(z)\,\mathrm{d}z, \quad \text{a.s..} \tag{6.4.2}$$

令 I_A 定义为 Borel 集合 $A \subseteq \mathcal{Z}$ 的简单函数, 这个函数不是连续的, 但可由一系列连续函数逼近. 选 ϕ 为 I_A, 从最后的结论推得测度 μ 的密度为 ρ^∞, μ 由下式定义给出:

$$\mu(A) = \lim_{T\to\infty} \frac{1}{T}\int_0^T I_A(z(t))\,\mathrm{d}t, \tag{6.4.3}$$

从而 $\mu(\mathrm{d}z) = \rho^\infty(z)\,\mathrm{d}z$, 而且不变分布 μ 清晰地度量了随机积分方程的一个样本路径在所给集合上停留的时间比例, 就像常微分方程和 Markov 链一样.

注 6.17 尽管没有证明前面定理的结论, 但有如下备注:

(i) 事实上, 当扩散矩阵 Γ 一致正定且具有周期边界条件时, 通过强极大值原理 (例 7.15) 可知 $\mathcal{L}$ 的零空间包含常数.

(ii) 如果假设 ρ^∞ 严格为正, 也可以直接由 (6.3.13) 得到同样的事实.

(iii) 时间平均 (6.4.2) 的收敛性可进行如下量化: 在 $\mathbb{T}^d$ 上, 设 $\mathcal{L}$ 是随机微分方程 (6.1.1) 的生成子, 其具有周期边界条件, 并且满足定理 6.16 的假设. 设 Φ 为**Poisson 方程**的解,

$$\mathcal{L}\Phi = \phi - \overline{\phi},$$

其中 $\overline{\phi}$ 定义为关于随机积分方程不变测度的标量值函数 ϕ 的平均. 通过 Fredholm 变换 (定理 2.42), Poisson 方程具有依赖于常数的唯一解. 对 $\Phi(x(t))$ 使用 Itô 公式得

$$\Phi(z(t)) = \Phi(z(0)) + \int_0^t \mathcal{L}\Phi(z(s))\,\mathrm{d}s + \int_0^t \langle \nabla\Phi(z(s)), \gamma(z(s))\,\mathrm{d}W(s)\rangle,$$

则

$$\frac{1}{T}(\Phi(z(T)) - \Phi(z(0))) = \frac{1}{T}\int_0^T \phi(z(s))\,\mathrm{d}s - \overline{\phi} + M_T,$$

$$M_T := \frac{1}{T}\int_0^T \langle \nabla \Phi(z(s)), \gamma(z(s))\,\mathrm{d}W(s)\rangle.$$

由 Itô 等距和关于 Φ 的假设有 $\mathbb{E}|M_T|^2 = O(1/T)$. 因此, 在 L^2 上有

$$\frac{1}{T}\int_0^T \Phi(z(s))\,\mathrm{d}s = \overline{\phi} + O\left(\frac{1}{\sqrt{T}}\right).$$

这本质上是遍历性定理, 描述了遍历性随机积分方程时间平均的收敛性①. 还注意到可通过鞅中心极限定理 (定理 3.33) 表示对该定理的修正, 由鞅中心极限定理表明 $\sqrt{T}M_T$ 收敛于 Brownian 运动, 该运动的协方差可用与生成子 $\mathcal{L}$ 有关的 Dirichlet 形式表示, 用 Poisson 方程 Φ 的解求值. □

例 6.18 在 $\mathbb{T}$ 上考虑一维 Brownian 运动

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \sigma\frac{\mathrm{d}W}{\mathrm{d}t}, \quad z(0) = z_0.$$

生成子 $\mathcal{L}$ 是微分算子

$$\mathcal{L} = \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}z^2},$$

该算子在 $[0,1]$ 上具有周期条件且是自共轭的. $\mathcal{L}$ 和 $\mathcal{L}^*$ 的零空间在 $[0,1]$ 上包含常数函数. 向后 Kolmogorov 方程和 Fokker-Planck 方程约化为在 $[0,1]$ 上具有周期边界条件的热传导方程

$$\frac{\partial\rho}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2\rho}{\partial x^2}.$$

Fourier 分析的直接应用表明解依指数收敛于一个常数, 见练习第 7 题. □

注意到当 Brownian 运动在环面上具有周期性时, 其在 $\mathbb{R}$ 上是不具有遍历性的, 就像 Lebesgue 测度作为一个概率测度, 不能够规范化一样. 然而, 尽管定理 6.16 适用于状态空间是 $\mathbb{T}^d$ 的情况, 通过为 $\mathcal{L}$ 和 $\mathcal{L}^*$ 选用合适的函数空间, 可以把定理 6.16 很容易推广到其他背景下. 下面在 $\mathbb{R}^d$ 中, 再举两个例子来说明这个问题.

例 6.19 考虑例 6.4 中的 OU 过程:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -\alpha z + \sqrt{2\lambda}\frac{\mathrm{d}W}{\mathrm{d}t}, \quad z(0) = z_0, \tag{6.4.4}$$

其中 $z(t) \in \mathbb{R}$. 假定 z_0 是固定的并且是非随机的, 应用 Itô 公式得到解

$$z(t) = \mathrm{e}^{-\alpha t}z_0 + \sqrt{2\lambda}\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}W(s),$$

则

$$\mathbb{E}z(t) = z_0\mathrm{e}^{-\alpha t},$$

① 它也是大数定理的一种形式.

并且由 Itô 等距同构得

$$\mathbb{E}\left(z\left(t\right)-\mathbb{E}z\left(t\right)\right)^{2}=2\lambda\mathbb{E}\left(\int_{0}^{t}\mathrm{e}^{-\alpha(t-s)}\mathrm{d}W\left(s\right)\right)^{2}=2\lambda\int_{0}^{t}\mathrm{e}^{-2\alpha(t-s)}\mathrm{d}s=\frac{\lambda}{\alpha}\left(1-\mathrm{e}^{-2\alpha t}\right).$$

由于 OU 过程 $z(t)$ 是 Gaussian 的, 因此, 前面的计算表明

$$z\left(t\right)\sim\mathcal{N}\left(m\left(t\right),\sigma^{2}\left(t\right)\right), \tag{6.4.5}$$

其中

$$m\left(t\right)=\mathrm{e}^{-\alpha t}z_{0},\quad \sigma^{2}\left(t\right)=\frac{\lambda}{\alpha}\left(1-\mathrm{e}^{-2\alpha t}\right),$$

表明当 $t\to\infty$ 时, 其收敛于 Gaussian 不变测度 $\mathcal{N}(0,\lambda/\alpha)$. 这是遍历性的表现形式.

由 (6.3.6) 得到 OU 过程的 Fokker-Planck 方程

$$\frac{\partial\rho}{\partial t}=\frac{\partial}{\partial t}\left(\alpha x\rho\right)+\lambda\frac{\partial^{2}\rho}{\partial x^{2}},$$

其初始条件是一个中心在 z_0 的 Dirac 质量. 容易证明其密度与 $z(t)$ 的 Gaussian 测度有关, 即

$$\rho\left(x,t\right)=\sqrt{\frac{1}{2\pi\sigma^{2}\left(t\right)}}\exp\left(-\frac{\left(x-m\left(t\right)\right)^{2}}{\sigma^{2}\left(t\right)}\right)$$

是这个线性偏微分方程的解, 从而由 $t\to\infty$ 时 Gaussian 随机变量 $\mathcal{N}(0,\lambda/\alpha)$ 的密度 ρ 收敛于 ρ^{∞}, 即 OU 过程的唯一不变的分布是

$$\rho^{\infty}\left(x\right)=\sqrt{\frac{\alpha}{2\lambda\pi}}\exp\left(-\frac{\alpha x^{2}}{2\lambda}\right). \tag{6.4.6}$$

□

例 6.20　考虑方程

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=-ax+y, & x\left(0\right)=x_{0},\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=-y+\sqrt{2\sigma}\dfrac{\mathrm{d}W}{\mathrm{d}t}, & y\left(0\right)=y_{0},\end{cases}$$

其中 $(x(t),y(t))\in\mathbb{R}^2$. 假定 $a>0$, Fokker-Planck 方程的形式为

$$\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x}\left(\left(-ax+y\right)\rho\right)+\frac{\partial}{\partial y}\left(-y\rho\right)=\sigma\frac{\partial^{2}\rho}{\partial y^{2}}.$$

前面的例子表明 y 具有 Gaussian 分布且收敛于一个 Gaussian 不变测度. 因为

$$x\left(t\right)=\mathrm{e}^{-at}x\left(0\right)+\int_{0}^{t}\mathrm{e}^{-a(t-s)}y\left(s\right)\mathrm{d}s,$$

并且 y 是 Gaussian 的, 于是推得 x 也是 Gaussian 的, 即由下式给出 e^{-at} 以及方差:

$$\mathbb{E}\left(\int_0^t \mathrm{e}^{-a(t-s)}y(s)\mathrm{d}s\right)^2.$$

这些想法表明找到了一个 Fokker-Planck 方程的稳态解, 形式如下:

$$\rho^{\infty}(x,y) \propto \exp\left(-\alpha x^2 + \beta xy - \gamma y^2\right).$$

(选取比例常数, 使得 ρ^{∞} 在 $\mathbb{R}^2$ 上积分为 1.) 代入得

$$\alpha = \frac{a(a+1)^2}{2\sigma}, \quad \beta = \frac{2a(a+1)}{2\sigma}, \quad \gamma = \frac{(a+1)}{2\sigma}.$$

注意已经找到关于 (x,y) 的 Gaussian 不变测度的密度

$$\rho(x,y) = \frac{1}{Z}\exp\left(-\frac{a(a+1)^2}{2\sigma}x^2 + \frac{2a(a+1)}{2\sigma}xy - \frac{(a+1)}{2\sigma}y^2\right),$$

其中 Z 为归一化常数. □

例 6.21 考虑随机微分方程

$$\frac{\mathrm{d}z}{\mathrm{d}t} = v(z) + \sqrt{2\sigma}\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中将 z 看成是 $\mathbb{T}^d$ 的一个元素且 $\sigma > 0$. 如果 v 为无散度向量值且是光滑的, 则唯一不变测度是 Lebesgue 测度. 注意到 Lebesgue 测度是不变的, 表明 $\mathcal{L}^*$ 是随机微分方程的生成子, 其中用 $-v(z)$ 替换 $v(z)$. 定理 6.16 保证了不变测度的唯一性. 而如果 $\sigma = 0$, 则 Lebesgue 测度为不变量, 此时不一定具有唯一性 (见 14 章). □

例 6.22 考虑随机积分

$$I(t) = \int_0^t \eta(s)\,\mathrm{d}W_1(s),$$

其中 $\eta(t)$ 为例 6.19 中定义的 OU 过程, 即

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\eta + \sqrt{2\sigma}\frac{\mathrm{d}W_2}{\mathrm{d}t}, \quad \eta(0) = \eta_0,$$

$W_1(t)$ 和 $W_2(t)$ 是独立的 Brownian 运动. η 的不变测度是一个服从分布 $\mathcal{N}(0,\sigma)$ 的 Gaussian 随机变量. 假定初始条件依不变测度分布, 因此, $\eta(t)$ 是一个平稳遍历 Markov 过程, 并且运用鞅中心极限定理 (推论 3.34) 得到

$$\lim_{\varepsilon\to 0}\varepsilon I\left(\frac{t}{\varepsilon^2}\right) = \sqrt{\sigma}W(t),$$

其中 $W(t)$ 是一个一维标准 Brown 运动. □

6.5 讨论和参考

有关随机微分方程的标准教科书见文献 [236, 12, 238, 114]; 用 Fokker-Planck 方程的观点讨论随机微分方程见文献 [117, 144, 218, 271, 320]; 对生成子 $\mathcal{L}$ 以及向后 Kolmogorov 方程的讨论见文献 [236]; 有关于椭圆性、次椭圆性以及方程解的光滑性的论述见文献 [275, 276]; 文献 [139] 很好地讨论了遍历性; 文献 [210] 概述了稳定性理论和随机微分方程的长时间特性; 在一些数学文献中, 有时称 Fokker-Planck 方程为向前 Kolmogorov 方程; 文献 [64] 概述了在应用数学中随机方法的应用.

只讨论了随机微分方程的强解, 弱解及其存在唯一性理论可在 [276,ch.5] 中找到. 随机微分方程弱形式等价于鞅形式, 见文献 [304].

在实际应用中有一些有趣的话题, 其中包括随机微分方程 (常微分方程) 与 Markov 链的耦合. 考虑随机微分方程 (6.1.1), 其中参数 $u \in \mathcal{I} = \{1, \cdots, n\}$,

$$\frac{\mathrm{d}z}{\mathrm{d}t} = h(z, u) + \gamma(z, u)\frac{\mathrm{d}W}{\mathrm{d}t}, \quad z(0) = z_0. \tag{6.5.1}$$

对于每个固定的 u, 令 $\mathcal{L}^u$ 定义为随机微分方程的生成子. 如果假定 u 由生成子为 $L(z)$ 的连续时间 Markov 链决定, 使得转移速率依赖于状态 z, 则对偶 (z, u) 形成一个 Markov 转换过程. 此过程实际就是 Markov 过程, 其概述见 [211]; 注意到对于每个固定的 z, $L(z) \in \mathbb{R}^{n\times n}$, 组合过程 (z, x) 的生成子形式如下:

$$\mathcal{L} = \begin{pmatrix} \mathcal{L}^1 & & & \\ & \mathcal{L}^2 & & \\ & & \ddots & \\ & & & \mathcal{L}^n \end{pmatrix} + L(z). \tag{6.5.2}$$

这个生成子是一个形式为 $v(z, u) = (v(z, 1), \cdots, v(z, u))^{\mathrm{T}}$ 的函数, 则 $v(z, u, t) = \mathbb{E}(\phi(z(t)), u(t) | z(t) = z, u(t) = u)$ 为向后 Kolmogorov 方程的解,

$$\frac{\partial v}{\partial t} = \mathcal{L}v,$$

其中 $v(z, u, 0) = \phi(z, u)$. 类似的概率密度向量 $\rho(z, u, t)$ 由下面的伴随方程生成:

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^* v,$$

其中 $\rho(z, u, 0)$ 描述了系统的初始概率. 由 Markov 链得到的随机微分方程, 其应用见 [326].

可通过不同的技巧研究随机微分方程的遍历特性, 包括泛函分析的偏微分方程的技巧 (见文献 [58, 196, 337]) 和概率技巧 (见文献 [139, 192, 227]). Lyapunov 函数对于证明随机微分方程遍历性是很有用的 [188, sec11.9].

本书中研究了有限维空间上的随机微分方程. 在无限维空间上还有一个成熟的随机微分方程的理论, 如随机偏微分方程. 在这种背景下, 基于半群的框架和遍历性的简要讨论见文献 [263].

6.6 练　习

1. 由 Stratonovich 随机微分方程 (6.2.1) 推导 Itô 随机微分方程 (6.2.3). 使用 Stratonovich 随机微分方程的 Itô 形式, 找到 Stratonovich 随机微分方程的 Fokker-Planck 方程.

2. 证明定理 6.6.(提示: 使用 Itô 公式以及随机积分的鞅性质.)

3. 考虑方程 (6.4.4) 中的 OU 过程:

a) 计算过程 $z(t)$ 的所有矩;

b) 验证本例中给出的 Fokker-Planck 方程解的计算公式, 证明 $\rho(0,x)=\delta(z-x)$;

c) 由前面的公式推导 OU 过程的长时间特性.

4. 考虑方程 (6.4.4) 中定义的 OU 过程, 证明生成子在由内积 (6.3.12) 定义的加权 L^2 空间上是自伴随的.

5. 考虑随机微分方程

$$m\ddot{x}=-\nabla V(x)-\gamma\dot{x}+\sqrt{\gamma D}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{6.6.1}$$

其中 m, D, γ 是正常数, 并且 $V(x):\mathbb{R}^d\mapsto\mathbb{R}$ 是一个光滑函数.

a) 写出方程 (6.6.1) 如下形式的随机微分方程的一阶系统:

$$\frac{\mathrm{d}z}{\mathrm{d}t}=-\frac{1}{D}K\nabla H(z)+\frac{1}{\sqrt{m}}J\nabla H(z)+\sqrt{K}\frac{\mathrm{d}B}{\mathrm{d}t},$$

其中 $z=\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}$, J(或 K) 是反对称 (或对称的) 矩阵, 定义 $y=\sqrt{m}\dot{x}$ 以及 $H(z)=1/2|y|^2+V(x)$;

b) 写出对应的生成子以及 Fokker-Planck 方程;

c) 解稳态 Fokker-Planck 方程 (提示: 使用变量分离);

d) 在一维情形下解方程 (6.6.1), 其中 $V(x)\equiv 0$, 并且 $V(x)=1/2x^2$.

6. 考虑 Markov 链 u, 其生成子为

$$L=\begin{pmatrix}-a & a\\ b & -b\end{pmatrix}.$$

现在 x 为一个随机微分方程的解, 其系数依赖于 u,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,u) + \alpha(x,u)\frac{\mathrm{d}W}{\mathrm{d}t}.$$

写出 Markov 过程 (x,u) 的生成子.

7. 在例 6.18 中用到的单位圆上考虑一个 Brownian 运动. 使用 Fourier 分析证明过程具有遍历性, 并且其稳定指数快速增长.

第 7 章　偏微分方程

7.1　格　　局

鉴于本书的要求, 本章概述椭圆型、抛物型以及双曲型偏微分方程的重要理论. 7.2 节研究椭圆问题, 以散度形式对椭圆型偏微分方程的 Dirichlet 周期边界问题开始, 介绍弱解的概念, 并运用 Lax-Milgram 定理证明了这类问题解的存在性和唯一性, 然后推导出椭圆型偏微分方程的 Fredholm 变换, 最后通过无散型算子的极大值原理研究了椭圆型偏微分方程; 7.3 节中研究了抛物型偏微分方程, 首先是说明弱解概念是怎样被推广到抛物型问题上的, 尤其是在有界空间域的问题上, 然后介绍椭圆型算子的极大值原理, 最后以无界域上的椭圆型偏微分方程的极大值原理结束; 7.4 节描述了有界域上双曲型偏微分方程的 Cauchy 问题; 7.5 节包括一个对半群的简短讨论, 以及如何统一对瞬态偏微分方程某一方面的理解; 7.6 节中提出了基本理论的各种推广, 并且作了一些书目备注.

第 2 章对本书所用的函数空间背景进行了描述. 使用古典解, 使得偏微分方程在区域内的每一点都成立, 从而在某种意义下所有要求的导数都存在且都连续, 使得方程具有相容性且满足边值条件; 接下来函数的推广 (可能是导数) 使得形成这一个区域的闭集; 有时称这一类古典解为强解. 本章也介绍了弱解和中性解的不同记号, 而相对来说, 这一类解缺少部分的光滑性.

7.2　椭圆型偏微分方程

Dirichlet 问题(均匀化)[①] 就是找到一个定义在开区间集合 $\Omega \in \mathbb{R}^d$ 上的函数 u, 求解

$$-\nabla \cdot (A\nabla u) = f, \quad x \in \Omega, \tag{7.2.1a}$$

$$u = 0, \quad x \in \partial\Omega, \tag{7.2.1b}$$

其中, $A = A(x)$ 是一个正定矩阵且 $f = f(x) \in H^{-1}(\Omega)$.

回顾前面的定义可知 H 为 $H^1_{\text{per}}(\mathbb{T}^d)$ 的子集, 见方程 (2.4.6). 周期问题就是找到一个 1 周期函数 u, 满足

$$-\nabla \cdot (A\nabla u) = f, \tag{7.2.2}$$

① 这里涉及边界条件的均匀化.

其中 $A = A(x)$ 是 1 周期正定矩阵且 $f = f(x) \in H^*$, 即 H 的对偶. 把这个问题看成圆环面 $\mathbb{T}^d$ 上的一个偏微分方程, 由第 2 章可知, 除非 f 在单位体上的积分为零, 否则, 无解. 由 (2.4.9) 给出的空间 H^* 保证了这个条件.

如下定义给出了 $A(x)$ 系数类:

定义 7.1 设 $\alpha, \beta \in \mathbb{R}$, 使得 $0 < \alpha \leqslant \beta < \infty$, 定义 $M(\alpha, \beta, \Omega)$ 为 $d \times d$ 矩阵 $A \in L^\infty\left(\Omega; \mathbb{R}^{d\times d}\right)$ 的集合, 使得对于每一个向量 $\xi \in \mathbb{R}^d$ 和任一 $x \in \Omega$ 都有

(i) $\langle \xi, A(x)\xi \rangle \geqslant \alpha |\xi|^2$;

(ii) $|A(x)\xi| \leqslant \beta |\xi|$.

此外, 定义 $M_{\text{per}}\left(\alpha, \beta, \mathbb{T}^d\right)$ 是 $\mathbb{T}^d$ 上具有周期系数的 $M\left(\alpha, \beta, \mathbb{T}^d\right)$ 矩阵的集合.

在前面的叙述中, 研究了如下形式的椭圆型算子:

$$\mathcal{A} = -\nabla \cdot (A\nabla) + b \cdot \nabla + c. \tag{7.2.3}$$

接下来, 将研究 Dirichlet 问题、周期问题以及 Fredholm 性质. 研究极大值原理时, 椭圆型算子将有如下形式:

$$\mathcal{A} = -A : \nabla\nabla + b \cdot \nabla + c. \tag{7.2.4}$$

如果 $A \in M(\alpha, \beta, \Omega)$, 则 (7.2.3) 或 (7.2.4) 式中的算子 $\mathcal{A}$ 被称为是一致椭圆的. 而 (7.2.3) 形式的算子以及相应的偏微分方程则被称为散度型算子. (7.2.4) 形式的算子称为无散型算子. 注意到, 假设矩阵 $A \in C^1(\Omega; \mathbb{R}^{d\times d})$, 则由 (7.2.3) 和 (7.2.4) 所定义的两种形式可以相互转换, 见练习第 1 题.

7.2.1 Dirichlet 问题

为了给定一个解的精确含义, 就需要介绍双线性形式

$$a[\phi, \psi] = \int_\Omega \langle A(x)\nabla\phi(x), \nabla\psi(x)\rangle \,\mathrm{d}x, \tag{7.2.5}$$

其中 $\phi, \psi \in H_0^1(\Omega)$. 注意到

$$a[\phi, \psi] = (A\nabla\phi, \nabla\psi),$$

其中 $(\cdot,\cdot)$ 是 $L^2(\Omega)$ 中的标准内积. 使用记号 $\langle\cdot,\cdot\rangle_{H^{-1},H_0^1}$ 表示 $H_0^1(\Omega)$ 和其对偶 $H^{-1}(\Omega)$ 所形成的对偶对 (见第 2 章).

定义 7.2 如果存在 $u \in H_0^1(\Omega)$, 使得

$$a[u, v] = \langle f, v\rangle_{H^{-1}H_0^1}, \quad \forall v \in H_0^1(\Omega) \tag{7.2.6}$$

成立, 则称 u 是边值问题 (7.2.1) 的一个弱解.

用一个测试函数 $v \in H_0^1(\Omega)$ 乘以偏微分方程 (7.2.1), 并通过分部积分就导出了这个解的概念. 对于所有 $v \in H_0^1(\Omega)$, 方程 (7.2.6) 都成立, 便给出了弱解的概念. 如果 $f \in L^2(\Omega)$, 则 (7.2.6) 可以写成

$$a[u,v] = (f,v), \quad \forall v \in H_0^1(\Omega). \tag{7.2.7}$$

运用 Lax-Milgram 定理 2.40 可以证明当系数矩阵 $A(x)$ 满足定义 7.1 中的条件时, 上述方程的弱解存在且唯一.

定理 7.3 Dirichlet 问题 (7.2.1) 中如果 $A \in M(\alpha,\beta,\Omega)$ 且 $f \in H^{-1}(\Omega)$, 则有唯一的弱解 $u \in H_0^1(\Omega)$, 而且如下估计成立:

$$\|u\|_{H_0^1(\Omega)} \leqslant \frac{1}{\alpha}\|f\|_{H^{-1}(\Omega)}. \tag{7.2.8}$$

证明 首先要验证 Lax-Milgram 定理的条件. 先证椭圆性. 由矩阵 A 的正定性可得

$$a[u,u] = \int_\Omega \langle A\nabla u, \nabla u\rangle\,\mathrm{d}x \geqslant \alpha\int_\Omega |\nabla u|^2\,\mathrm{d}x = \alpha\|u\|_{H_0^1(\Omega)}^2.$$

接下来证连续性. 利用 L^∞ 在 A 上有界以及 Cauchy-Schwarz 不等式可估计

$$\begin{aligned} a[u,v] &= \int_\Omega \langle A\nabla u, \nabla v\rangle\,\mathrm{d}x \leqslant \beta\int_\Omega |\nabla u|\,|\nabla v|\,\mathrm{d}x \\ &\leqslant \beta\|\nabla u\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} = \beta\|u\|_{H_0^1(\Omega)}\|v\|_{H_0^1(\Omega)}. \end{aligned}$$

双线性形式 $a[u,v]$ 满足 Lax-Milgram 定理的条件, 因此, (7.2.6) 存在唯一解 $u \in H_0^1(\Omega)$.

最后证明估计 (7.2.8). 运用广义 Cauchy-Schwarz 不等式 (2.4.5) 可得

$$\alpha\|u\|_{H_0^1(\Omega)}^2 \leqslant a[u,u] = \langle f,u\rangle \leqslant \|f\|_{H^{-1}(\Omega)}\|u\|_{H_0^1(\Omega)},$$

由这个式子可得估计 (7.2.8).

注 7.4 如果问题 (7.2.1) 有 $A \in M(\alpha,\beta,\Omega)$ 以及 $f \in L^2(\Omega)$, 则

$$\|u\|_{H_0^1(\Omega)} \leqslant \frac{C_\Omega}{\alpha}\|f\|_{L^2(\Omega)},$$

其中 C_Ω 是定义在定理 2.21 中依赖于区域 Ω 的 Poincaré常数.

边界 (7.2.8) 能使获得更多 Dirichlet 问题单参数簇解的信息, 可以由 A 上参数的相关性推导得出.

定理 7.5 假设存在正系数 $\alpha,\beta\,(\alpha \leqslant \beta)$, 使得对于所有的 $\varepsilon > 0$, 矩阵 $A^\varepsilon = A^\varepsilon(x)$ 的单参数簇属于 $M(\alpha,\beta,\Omega)$. 考虑 Dirichlet 问题

$$-\nabla\cdot(A^\varepsilon\nabla u^\varepsilon) = f, \quad x \in \Omega, \tag{7.2.9a}$$

$$u^{\varepsilon}=0,\quad x\in\partial\Omega, \tag{7.2.9b}$$

并且 $f=f(x)\in H^{-1}(\Omega)$, 则存在一个独立于 ε 的常数 C, 使得

$$\|u^{\varepsilon}\|_{H_0^1(\Omega)}\leqslant C; \tag{7.2.10}$$

此外, 存在一个子序列 $\{\varepsilon_n\}_{n\geqslant 0}$ 和一个函数 $u\in H_0^1(\Omega)$, 使得在 $L^2(\Omega)$ 上有如下强收敛:

$$u^{\varepsilon_n}(x)\to u(x).$$

证明　(7.2.8) 暗含 (7.2.10). 定理 2.1.9 中的 Rellich 紧性定理暗含了存在一个函数 $u\in H_0^1(\Omega)$ 和一个子序列 $\{\varepsilon_n\}\in\varepsilon$, 使得 u 是 u^{ε_n} 的强 L^2 极限. □

注 7.6　第 12 章和第 19 章研究椭圆型偏微分方程的均匀化时, 将研究 (7.2.9), 并且找出满足极限 u 的方程. □

7.2.2　周期问题

周期问题的解有可能仅由一个常数决定, 这很直观明了. 固定这个常数就可以确定解的唯一性, 这就是研究 H 的原因. 用符号 $a_1[\cdot,\cdot]$ 表示双线性形式

$$a_1[u,v]=\int_{\mathbb{T}^d}\langle A(x)\nabla u(x),\nabla v(x)\rangle\,\mathrm{d}x,\quad \forall u,v\in H. \tag{7.2.11}$$

用 $\langle\cdot,\cdot\rangle_{H^*,H}$ 表示 H 和 H^* 之间的对偶对 (见第 2 章).

定义 7.7　如果

$$a_1[u,v]=\langle f,v\rangle_{H^*,H},\quad \forall v\in H, \tag{7.2.12}$$

则称 $u\in H$ 是边值问题的一个弱解.

H^* 的构造意味着当 $f\in H^*$ 时, (7.2.25) 有唯一解, 因为它保证了 f 有零均值. 因此, 在空间 H 上, (7.2.2) 的弱解存在且唯一.

定理 7.8　如果有 $A\in M_{\mathrm{per}}(\alpha,\beta,\mathbb{T}^d)$ 和 $f\in H^*$, 则问题 (7.2.2) 有唯一弱解 $u\in H$, 使得如下估计成立:

$$\|u\|_H\leqslant\frac{1}{\alpha}\|f\|_{H^*}. \tag{7.2.13}$$

定理 7.8 的证明基本上和定理 7.3 相同, 在此省略. 事实上, 仅当 f 有零均值时, (7.2.2) 才有唯一解, 这个条件由 $f\in H^*$ 确定, 也可以通过 Fredholm 性质 (接下来将要转移到的主题) 的方式表示.

7.2.3　Fredholm 性质

在本节中证明散度形式下的椭圆型微分算子

$$\mathcal{A}=-\nabla\cdot(A\nabla)+b\nabla+c, \tag{7.2.14}$$

其中周期系数 A, b, c 以及周期边界条件满足 Fredholm 性质. 注意到因为 $\mathcal{A}$ 是一个无界算子, 所以不能直接应用定理 2.42. 主要思想是针对如下预解算子:

$$R_{\mathcal{A}}(\lambda) = (\mathcal{A} + \lambda I)^{-1} \tag{7.2.15}$$

进行研究, 其中 I 代表 $L^2_{\mathrm{per}}(\mathbb{T}^d)$ 上的恒等算子, 而且 $\lambda > 0$. 将证明这个算子是紧的, 因为 λ 足够大, 所以可以运用 Fredholm 定理.

对算子 $\mathcal{A}$ 的系数作如下假设:

$$A \in M_{\mathrm{per}}(\alpha, \beta, \mathbb{T}^d), \tag{7.2.16a}$$

$$A(x) = A(x)^{\mathrm{T}}, \quad \forall x \in \overline{\Omega}, \tag{7.2.16b}$$

$$b \in C^1_{\mathrm{per}}(\mathbb{T}^d), \tag{7.2.16c}$$

$$c \in L^\infty_{\mathrm{per}}(\mathbb{T}^d). \tag{7.2.16d}$$

$\mathcal{A}$ 的 L^2 伴随矩阵 $\mathcal{A}^*$ 由下式给出:

$$\mathcal{A}^* U = -\nabla \cdot (A \nabla U) - \nabla \cdot (bU) + cU, \tag{7.2.17}$$

它也具有周期边界条件. 对于 $f, F \in L^2_{\mathrm{per}}(\mathbb{T}^d)$, 研究偏微分方程

$$\mathcal{A} u = f, \quad u\text{是 1 周期}, \tag{7.2.18}$$

以及它的伴随矩阵

$$\mathcal{A}^* U = F, \quad U\text{是 1 周期}. \tag{7.2.19}$$

设 $a[\cdot,\cdot]$, $a^*[\cdot,\cdot]$: $H^1_{\mathrm{per}}(\mathbb{T}^d) \times H^1_{\mathrm{per}}(\mathbb{T}^d) \to \mathbb{R}$ 表示与 $\mathcal{A}$ 和 $\mathcal{A}^*$ 相关的双线性形式, 分别为

$$a[u,v] = \int_{\mathbb{T}^d} (\langle A\nabla u, \nabla v\rangle + (b \cdot \nabla u) v + cuv)\, \mathrm{d}x, \quad \forall u, v \in H^1_{\mathrm{per}}(\mathbb{T}^d) \tag{7.2.20}$$

和

$$a^*[u,v] = \int_{\mathbb{T}^d} (\langle A\nabla u, \nabla v\rangle - \nabla \cdot (bu) v + cuv)\, \mathrm{d}x, \quad \forall u, v \in H^1_{\mathrm{per}}(\mathbb{T}^d).$$

正如前面小节中, 只要

$$a[u,v] = (f,v), \quad \forall v \in H^1_{\mathrm{per}}(\mathbb{T}^d) \tag{7.2.21}$$

和

$$a^*[U,V] = (F,V), \quad \forall V \in H^1_{\mathrm{per}}(\mathbb{T}^d) \tag{7.2.22}$$

成立, 则 u 和 U 分别是偏微分方程 (7.2.18) 和 (7.2.19) 的弱解.

用 $\mathcal{N}$ 表示一个算子的零空间. 下面的定理包含了这一主要结论.

定理 7.9(周期性椭圆型偏微分方程的 Fredholm 性质)　假设满足 (7.2.16) 中的条件, 则以下性质成立:

(i) 对于每一个 $f \in L^2_{\text{per}}(\mathbb{T}^d)$, (7.2.18) 存在唯一解; 或者

(ii) 均匀化方程

$$\mathcal{A}u = 0, \quad u\text{是 1 周期} \tag{7.2.23}$$

至少有一个非零解, 并且有

$$1 \leqslant \dim(\mathcal{N}(\mathcal{A})) = \dim(\mathcal{N}(\mathcal{A}^*)) < \infty.$$

在这种情况下, 当且仅当

$$(f, v) = 0, \quad \forall v \in \mathcal{N}(\mathcal{A}^*)$$

时, 有界值问题 (7.2.18) 有一个弱解.

定理 7.9 的证明将用到如下两个引理:

引理 7.10　假设满足条件 (7.2.16), 则存在常数 $\nu, \mu > 0$, 使得

$$|a[u, v]| \leqslant C\|u\|_{H^1}\|v\|_{H^1}$$

和

$$\frac{\alpha}{2}\|u\|_{H^1}^2 \leqslant a[u, u] + \mu\|u\|_{L^2}^2$$

对于所有的 $u, v \in H^1_{\text{per}}(\mathbb{T}^d)$ 成立.

证明　(1) 根据 L^∞ 在系数 A 和 b 上有界以及 Cauchy-Schwarz 不等式推导可得

$$\begin{aligned}
|a(u, v)| \leqslant & \left|\int_{\mathbb{T}^d}\langle A\nabla u, \nabla v\rangle \mathrm{d}x + \int_{\mathbb{T}^d}(b \cdot \nabla u)v\mathrm{d}x + \int_{\mathbb{T}^d} cuv\mathrm{d}x\right| \\
\leqslant & \|A\|_{L^\infty}\int_{\mathbb{T}^d}|\nabla u||\nabla v|\,\mathrm{d}x + \|b\|_{L^\infty}\int_{\mathbb{T}^d}|\nabla u||v|\,\mathrm{d}x + \|c\|_{L^\infty}\int_{\mathbb{T}^d}|u||v|\,\mathrm{d}x \\
\leqslant & C\left(\|\nabla u\|_{L^2}\|\nabla v\|_{L^2} + \|\nabla u\|_{L^2}\|v\|_{L^2} + \|u\|_{L^2}\|v\|_{L^2}\right) \\
\leqslant & C\|u\|_{H^1}\|v\|_{H^1}.
\end{aligned}$$

(2) 使用一致椭圆性计算

$$\begin{aligned}
\alpha\|\nabla u\|_{L^2}^2 \leqslant & \int_{\mathbb{T}^d}\langle \nabla u, A\nabla u\rangle \mathrm{d}x \\
= & a[u, u] - \int_{\mathbb{T}^d}(b\nabla u)u\mathrm{d}x - \int_{\mathbb{T}^d} c|u|^2\mathrm{d}x
\end{aligned}$$

$$\leqslant a[u,u]+\int_{\mathbb{T}^d}\left(\|b\|_{L^\infty}|\nabla u||u|+\|c\|_{L^\infty}|u|^2\right)\mathrm{d}x. \tag{7.2.24}$$

运用代数不等式

$$ab\leqslant\delta a^2+\frac{1}{4\delta}b^2,\quad \forall\delta>0,$$

把这个不等式代入方程 (7.2.2) 的右边可得

$$\int_{\mathbb{T}^d}|\nabla u||u|\,\mathrm{d}x\leqslant\delta\|\nabla u\|_{L^2}^2+\frac{1}{4\delta}\|u\|_{L^2}^2. \tag{7.2.25}$$

选择 δ, 使得

$$\alpha-\|b\|_{L^\infty}\delta=\frac{\alpha}{2}.$$

用不等式 (7.2.25) 和 (7.2.24) 中 δ 的选值可得

$$\frac{\alpha}{2}\|\nabla u\|_{L^2}^2\leqslant a[u,u]+\frac{1}{4\delta}\|b\|_{L^\infty}\|u\|_{L^2}^2+\|c\|_{L^\infty}\|u\|_{L^2}^2.$$

在前面的不等式两边均加上 $\frac{\alpha}{2}\|u\|_{L^2}^2$ 可得

$$\frac{\alpha}{2}\|u\|_{H^1}^2\leqslant a[u,u]+\mu\|u\|_{L^2}^2,$$

其中 $\mu=\frac{1}{4\delta}\|b\|_{L^\infty}+\|c\|_{L^\infty}+\frac{\alpha}{2}$. □

引理 7.11 假设条件 (7.2.16) 成立, 从引理 7.10 取 μ, 则对于每个 $\lambda\geqslant\mu$ 和函数 $f\in L^2_{\mathrm{per}}(\mathbb{T}^d)$, 问题

$$(\mathcal{A}+\lambda I)u=f,\quad u(x)\text{是 1 周期} \tag{7.2.26}$$

都有唯一弱解 $u\in H^1_{\mathrm{per}}(\mathbb{T}^d)$.

证明 设 $\lambda\geqslant\mu$, 定义算子

$$\mathcal{A}_\lambda:=\mathcal{A}+\lambda I. \tag{7.2.27}$$

$\mathcal{A}_\lambda$ 相关的双线性形式是

$$a_\lambda[u,v]=a[u,v]+\lambda(u,v),\quad \forall u,v\in H^1_{\mathrm{per}}(\mathbb{T}^d). \tag{7.2.28}$$

接下来由引理 7.10 及假设 $\lambda\geqslant\mu$ 可知, 双线性形式 $a_\lambda[u,v]$ 在 $H^1_{\mathrm{per}}(\mathbb{T}^d)$ 上连续且具有椭圆性. 因此, 由 Lax-Milgram 定理适用①可知方程

$$a_\lambda[u,v]=(g,v),\quad \forall v\in H^1_{\mathrm{per}}(\mathbb{T}^d) \tag{7.2.29}$$

① 设 $H=H^1_{\mathrm{per}}(\mathbb{T}^d)$. 由 $f\in L^2_{\mathrm{per}}(\mathbb{T}^d)$ 可知 $\langle f,v\rangle_{H^*,H}=(f,v)$, 并且这定义了 $H^1_{\mathrm{per}}(\mathbb{T}^d)$ 上的一个有界线性函数.

存在唯一解 $u \in H^1_{\text{per}}(\mathbb{T}^d)$. 边值问题 (7.2.26) 的弱形式是非常精确的. □

定理 7.9 的证明　(1) 由引理 7.11 知道对于每一个 $g \in L^2_{\text{per}}(\mathbb{T}^d)$, 方程

$$a_\mu[u,v] = (g,v), \quad \forall v \in H^1_{\text{per}}(\mathbb{T}^d) \tag{7.2.30}$$

都存在唯一解 $u \in H^1_{\text{per}}(\mathbb{T}^d)$. 用 (7.2.15) 中定义的预解算子表示以下形式:

$$u = R_{\mathcal{A}}(\mu)\, g \tag{7.2.31}$$

的解. 考虑 (7.2.18), 在其两边均加上 μu 可得

$$\mathcal{A}_\mu u = \mu u + f,$$

其中 $\mathcal{A}_\mu$ 如 (7.2.27) 中所定义. 这个方程的弱形式是

$$a_\mu[u,v] = (\mu u + f, v), \quad \forall v \in H^1_{\text{per}}(\mathbb{T}^d).$$

也可以将其重写为一个积分方程 (见 (7.2.31))

$$u = R_{\mathcal{A}}(\mu)(\mu u + f),$$

或者等价于

$$(I-K)\,u = h,$$

其中

$$K := \mu R_{\mathcal{A}}(\mu), \quad h = R_{\mathcal{A}}(\mu)\, f.$$

(2) 接着讨论算子 $K: L^2_{\text{per}}(\mathbb{T}^d) \to L^2_{\text{per}}(\mathbb{T}^d)$ 是紧的. 确实, 设 u 是 (7.2.30) 的一个解, 这是由 (7.2.31) 给出的. 运用引理 7.10 中的第二个假设, 由双线性形式 (7.2.28) 以及 (7.2.30) 中的 Cauchy-Schwarz 不等式可得

$$\frac{\alpha}{2}\|u\|^2_{H^1} \leqslant a_\mu[u,u] = (g,u) \leqslant \|g\|_{L^2}\|u\|_{L^2} \leqslant \|g\|_{L^2}\|u\|_{H^1},$$

所以有

$$\|u\|_{H^1} \leqslant \frac{2}{\alpha}\|g\|_{L^2}.$$

由 (7.2.31), K 的定义以及之前的估计推导可得

$$\|Kg\|_{H^1} \leqslant \mu\|u\|_{H^1} \leqslant \frac{2}{\alpha}\mu\|g\|_{L^2}. \tag{7.2.32}$$

由 Rellich 紧性定理知 $H^1_{\text{per}}(\mathbb{T}^d)$ 是紧嵌入到 $L^2_{\text{per}}(\mathbb{T}^d)$ 中的, 因此, (7.2.32) 估计暗含了 $H^1_{\text{per}}(\mathbb{T}^d)$ 中的有界集是 $L^2_{\text{per}}(\mathbb{T}^d)$ 中的紧集. 因此, 它是一个紧性算子.

(3) 将 Fredholm 性质 (定理 2.42) 应用到算子 K, 则有

(a) 存在唯一的 $u \in L^2_{\text{per}}(\mathbb{T}^d)$, 使得

$$(I-K)u = h \tag{7.2.33}$$

或者

(b) 齐次方程

$$(I-K)u = 0 \tag{7.2.34}$$

存在一个非零解 $u \in L^2_{\text{per}}(\mathbb{T}^d)$. 假设 (a) 成立, 给出 (7.2.33) 的唯一解, 则从前面的分析中可以推出 (7.2.18) 存在唯一弱解 $u \in H^1_{\text{per}}(\mathbb{T}^d)$. 若假设 (b) 成立, 那么 (7.2.34) 有一个非零解. 设 N 和 N^* 分别表示零空间 $I-K$ 和 $I-K^*$ 的维数. 由定理 2.42 知 $N=N^*$, 此外, 也可以直接证明

$$u \in \mathcal{N}(I-K) = 0 \Leftrightarrow a[u,\phi] = 0, \quad \forall u \in H^1_{\text{per}}(\mathbb{T}^d)$$

和

$$u \in \mathcal{N}(I-K^*) = 0 \Leftrightarrow a^*[v,\phi] = 0, \quad \forall \phi \in H^1_{\text{per}}(\mathbb{T}^d).$$

因此, K 的 Fredholm 性质暗含了 $\mathcal{A}$ 的 Fredholm 性质 (在弱解的背景下).

(4) 下面证明定理的最后一部分. 设 $v \in \mathcal{N}(I-K^*)$, 由定理 2.42 可知, (7.2.34) 有解的充要条件是

$$(h,v) = 0, \quad \forall v \in \mathcal{N}(I-K^*).$$

由于

$$(h,v) = (R_{\mathcal{A}}(\mu)f, v) = \frac{1}{\mu}(Kf,v) = \frac{1}{\mu}(f,K^*v) = \frac{1}{\mu}(f,v).$$

因此, 问题 (7.2.18) 有一个弱解的充分必要条件为

$$(f,v) = 0, \quad \forall v \in \mathcal{N}(\mathcal{A}^*).$$

这就完成了定理的证明.

例 7.12 设 $f \in L^2_{\text{per}}(\mathbb{T}^d)$ 且 $A(x)$ 满足假设 (7.2.16a) 和 (7.2.16b), 则问题

$$a_1[u,\phi] = (f,\phi), \quad \forall \phi \in H^1_{\text{per}}(\mathbb{T}^d)$$

有唯一解 $u \in H$ 的充分必要条件为

$$(f,1) = 0, \tag{7.2.35}$$

其中 $a_1[\cdot,\cdot]$ 的定义见 (7.2.11). 实际上, 考虑均匀化伴随矩阵方程

$$\mathcal{A}^*v = 0.$$

显然, 常数函数 ($v=1$) 是这个方程的一个解, 矩阵 $A(x)$ 的一致椭圆性暗含了

$$\int_{\mathbb{T}^d} |\nabla v|^2 \mathrm{d}x = 0,$$

相对于 Lebesgue 测度来说, v 是一个常数, 因此, 常数解是唯一的. 由于满足假设条件 (7.2.16a) 和 (7.2.16b), 则定理 7.9 适用且结论成立.

注 7.13　在 Fredholm 性质和其他背景中, 使用恒等式

$$\int_{\mathbb{T}^d} \nabla \cdot q(y) \mathrm{d}y = 0,$$

对于任意 1 周期 C^1 函数 q 都成立. 这也遵循散度定理. 通过单位正立方体 (环面)$\mathbb{T}^d$, 并运用周期性可知 q 的总流量是零. □

7.2.4　极大值原理

本小节考虑无散度形式中椭圆微分算子的极大值问题. 特别地, 考虑 (7.2.4) 中的算子 $\mathcal{L}$,

$$\mathcal{L} = A(x) : \nabla\nabla - b(x) \cdot \nabla - c(x), \tag{7.2.36}$$

其中假设 $A, b, c \in C(\overline{\Omega})$ 且 $A \in M(\alpha, \beta, \Omega)$, 则算子 $\mathcal{L}$ 在 Ω 上一致椭圆, 使用符号

$$u^+ = \max\{u, 0\}, \quad u^- = -\min\{u, 0\}. \tag{7.2.37}$$

假设 $\partial\Omega = \overline{\Omega} \backslash \Omega$ 是光滑的, 通常用 $\partial\Omega$ 表示区域 Ω 上的边界. 首先来说明无散度形式中一致椭圆算子的强极大值原理.

定理 7.14　设 $u \in C^2(\Omega) \bigcap C(\overline{\Omega})$ 和 $c: \Omega \to \mathbb{R}^+$, 并设 $\mathcal{L}$ 是无散度形式的椭圆算子 (7.2.36).

(i) 如果 Ω 上有 $\mathcal{L}u \geqslant 0$, 则

$$\max_{\overline{\Omega}} u \leqslant \max_{\partial\Omega} u^+;$$

(ii) 如果 Ω 上有 $\mathcal{L}u \leqslant 0$, 则

$$\min_{\overline{\Omega}} u \geqslant -\max_{\partial\Omega} u^-.$$

如果 Ω 上有 $\mathcal{L}u = 0$, 则

$$\max_{\overline{\Omega}} |u| = \max_{\partial\Omega} |u|.$$

例 7.15　在 $\mathcal{Z} = \mathbb{T}^d$ 的情况下, 考虑偏微分方程 (6.1.2). 生成子 $\mathcal{L}$ 是由 (6.3.2) 给出的周期边界条件. 假设 $\Gamma \in M_{\mathrm{per}}(\alpha, \beta, \mathbb{T}^d)$ 是这个生成子为零的特征值相对应

的特征函数, 在 $\Omega_0=[0,1]^d$ 上满足 $\mathcal{L}\psi=0$. 因此, 对于任意这样的函数 ψ, $|\psi|$ 的极大值都能达到 $[0,1]^d$ 上的边界值. 然而, 由于 ψ 是定义在 $\mathbb{T}^d$ 上的, 因此, 可以推导出对于任意 $a\in\mathbb{R}^d$, 在 $\Omega_a=a+[0,1]^d$ 有 $\mathcal{L}\psi=0$. 对于任意 $a\in\mathbb{R}^d$, $|\psi|$ 在边界 Ω_a 上达到最大值, 所以 $|\psi|$ 必须是一个常数函数. 结合 ψ 的连续性可以推导出 $\mathcal{L}$ 在 $\mathbb{T}^d$ 上由常数所张成的一维零空间. 由 Fredholm 性质可知 $\mathcal{L}^*$ 也有一个由 ρ^∞ 所张成一维零空间. 事实上, 零空间是一维的, 可以由随机过程的遍历性体现出来. □

运用极大值原理可以获得椭圆型偏微分方程解的先验界, 这种形式的结论如下:

推论 7.16 设 Ω 是 $\mathbb{R}^d$ 中的一个有界区域且为光滑边界, 又设 $u\in C(\overline{\Omega})\bigcap C^2(\Omega)$ 是如下非均匀化 Dirichlet 问题的一个古典解:

$$\begin{aligned}\mathcal{L}u&=f,\quad x\in\Omega,\\ u&=g,\quad x\in\partial\Omega,\end{aligned}$$

其中 $\mathcal{L}$ 由 (7.2.36) 给出且系数光滑, A 使得 $-\mathcal{L}$ 具有一致椭圆性, b 有界, $c\geqslant 0$, 则称解 u 满足如下估计:

$$\|u\|_{L^\infty(\Omega)}\leqslant\|g\|_{L^\infty(\partial\Omega)}+C\|f\|_{L^\infty(\Omega)}. \tag{7.2.38}$$

常数 C 取决于 $\|b\|_{L^\infty(\Omega)}$ 上区域的大小以及一致椭圆性常数 α.

7.3 抛物型偏微分方程

7.3.1 有界域

假设 $\mathcal{L}$ 为离散形式下的一个椭圆型微分算子, 由 (7.2.3) 给出. 设 $\Omega\in\mathbb{R}^d$ 为有界开集, $T>0$, 并且定义 $\Omega_T=\Omega\times(0,T)$. 考虑如下抛物型偏微分方程:

$$\frac{\partial u}{\partial t}=\mathcal{L}u+f,\quad (x,t)\in\Omega_T, \tag{7.3.1a}$$

$$u=0,\quad (x,t)\in\partial\Omega\times(0,T], \tag{7.3.1b}$$

$$u=g,\quad (x,t)\in\overline{\Omega}\times\{0\}. \tag{7.3.1c}$$

假设 $A,b,c\in L^\infty(\Omega)$, $f\in L^2(\Omega_T)$, 并且 $g\in L^2(\Omega)$, 也可以假设 $A\in M(\alpha,\beta,\Omega)$, 使得 $-\mathcal{L}$ 为一致椭圆性.

如 (7.2.20) 定义双线性形式 $a[u,v]$. 将解 u 看成一个函数 $u:[0,T]\to H_0^1(\Omega)$, 并和弱解的以下定义起作用.

定义 7.17 $u\in L^2((0,T);H_0^1(\Omega))$ 是 (7.3.1) 的一个弱解, 必须满足

- $\dfrac{\mathrm{d}u}{\mathrm{d}t} \in L^2\left((0,T);H^{-1}(\Omega)\right)$;
- $\left\langle \dfrac{\mathrm{d}u}{\mathrm{d}t}, v\right\rangle_{H^{-1},H_0^1} + a[u,v] = (f,v), \forall v \in H_0^1(\Omega), \text{a.e. } t \in (0,T)$;
- $u(0) = g$.

注 7.18 Sobolev 空间的标准特性暗含了 $u \in C\left([0,T];L^2\left(\mathbb{R}^d\right)\right)$, 因此, 也使得方程 $u(0) = g$ 有意义. □

定理 7.19 在所述假设成立的情况下, (7.3.1) 存在唯一弱解, 而且对于 $u' = \mathrm{d}u/\mathrm{d}t$, 此解满足

$$
\begin{aligned}
&\|u\|_{L^\infty((0,T);L^2(\Omega))} + \|u\|_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|u'\|_{L^2((0,T);H^{-1}(\Omega))} \\
\leqslant & C\left(\|f\|_{L^2((0,T);L^2(\Omega))} + \|g\|_{L^2(\Omega)}\right).
\end{aligned} \tag{7.3.2}
$$

方法论曾用能量法来证明定理 7.19. (7.3.2) 中的估计通常叫做能量估计. 弱解的运用也可能被拓展到无界域. 将在 7.4 节中的双曲方程中对这些进行说明.

7.3.2 极大值原理

极大值原理对于抛物型偏微分方程的研究是非常有用的, 并且通过它可以得到一个先验估计. 使用记号

$$
\Omega_T = \Omega \times (0,T],
$$

$$
\partial\Omega_T = \partial\Omega \times \{0\}.
$$

因此有 $\partial\Omega = \overline{\Omega}\backslash\Omega$ 和 $\partial\Omega_T = \overline{\Omega}_T\backslash\Omega_T$. 如在椭圆型的情况下, 假设 $A, b, c \in C(\Omega)$ 和 $A \in M(\alpha,\beta,\Omega)$ 使得算子 $\mathcal{L}$ 在 Ω 上一致椭圆, 如 (7.2.37) 使用记号

$$
u^+ = \max\{u,0\}, \quad u^- = -\min\{u,0\}.
$$

定理 7.20 设 $u \in C^{2,1}(\Omega_T)\bigcap C\left(\overline{\Omega}_T\right)$ 和 $c : \Omega \to \mathbb{R}^+$, 考虑无散形式下的椭圆型算子 $\mathcal{L}$.

(i) 如果在 Ω_T 上 $\partial u/\partial t - \mathcal{L}u \leqslant 0$, 则

$$
\max_{\overline{\Omega}_T} u \leqslant \max_{\partial\Omega_T} u^+;
$$

(ii) 如果在 Ω_T 上 $\partial u/\partial t - \mathcal{L}u \geqslant 0$, 则

$$
\min_{\overline{\Omega}_T} u \geqslant -\max_{\partial\Omega_T} u^-,
$$

那么如果在 Ω_T 上 $\partial u/\partial t - \mathcal{L}u = 0$, 则

$$
\max_{\overline{\Omega}_T} |u| = \max_{\partial\Omega_T} |u|.
$$

抛物型偏微分方程强极大值原理的一个重要推论就是如下的先验估计.

推论 7.21 设 $\varOmega$ 是 $\mathbb{R}^d$ 上的一个有界区域且有光滑边界, 并设 $u \in C^{2,1}(\varOmega_T,\mathbb{R}) \bigcap C(\overline{\varOmega}_T,\mathbb{R})$ 是如下方程的一个古典解:

$$\frac{\partial u}{\partial t} - \mathcal{L}u = f, \quad (x,t) \in \varOmega \times \mathbb{R}^+,$$

其中 $u(x,0) = g(x)$ 和 $c(x) \equiv 0$, $\mathcal{L}$ 由 (7.2.36) 给出且系数光滑, A 使得 $-\mathcal{L}$ 一致椭圆性, b 有界, $c \geqslant 0$. 假设在 $\partial\varOmega$ 上 $u=0$, 则 u 满足如下估计:

$$\max_{\varOmega_T} |u| \leqslant \|g\|_{L^\infty(\varOmega)} + \int_0^T \|f(\cdot,s)\|_{L^\infty(\varOmega)} \mathrm{d}s. \tag{7.3.3}$$

证明 设

$$v(t) = \|g\|_{L^\infty(\varOmega)} + \int_0^t \|f(\cdot,s)\|_{L^\infty(\varOmega)} \mathrm{d}s,$$

注意到

$$\frac{\partial v}{\partial t} - \mathcal{L}v = \|f(\cdot,t)\|_{L^\infty(\varOmega)},$$

并且 $v(0) = \|g\|_{L^\infty(\varOmega)}$, 则对于 $e = u - v$ 有

$$\frac{\partial e}{\partial t} - \mathcal{L}e \leqslant 0.$$

由于在 $\partial\varOmega_T$ 上 $e \leqslant 0$, 由极大值原理 (定理 7.20) 推导可得

$$\max_{\varOmega_T} e \leqslant 0,$$

因此, 在 $\overline{\varOmega_T}$ 中有 $u \leqslant v$, 则

$$\max_{\overline{\varOmega}_T} u \leqslant \|g\|_{L^\infty(\varOmega)} + \int_0^T \|f(\cdot,s)\|_{L^\infty(\varOmega)} \mathrm{d}s.$$

类似地, 对于 $d = u + v$ 有

$$\frac{\partial d}{\partial t} - \mathcal{L}d \geqslant 0.$$

由于在 $\partial\varOmega_T$ 上 $d \geqslant 0$, 由极大值原理 (定理 7.20) 推导可得

$$\min_{\overline{\varOmega}_T} d \geqslant 0.$$

因此, 在 $\overline{\varOmega}_T$ 上 $-u \leqslant v$, 则

$$\min_{\overline{\varOmega}_T} (-u) \leqslant \|g\|_{L^\infty(\varOmega)} + \int_0^T \|f(\cdot,s)\|_{L^\infty(\varOmega)} \mathrm{d}s.$$

结合上下界给出 (7.3.3). □

7.3.3　无界域：Cauchy 问题

在本节中，将描述无散形式下的抛物型偏微分方程的基本理论以及无界域. 特别要研究如下初值 (Cauchy) 问题：

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f, \quad (x,t) \in \mathbb{R}^d \times (0,T], \tag{7.3.4a}$$

$$u = g, \quad (x,t) \in \mathbb{R}^d \times \{0\}, \tag{7.3.4b}$$

其中

$$\mathcal{L} := b(x) \cdot \nabla + \frac{1}{2} A(x) : \nabla\nabla. \tag{7.3.5}$$

假设 $A, b \in C_b^{\infty}(\mathbb{R}^d)$, $f \in C_0^{\infty}(\mathbb{R}^d \times [0,T])$ 且 $A \in M(\alpha, \beta, \mathbb{R}^d)$, $0 < \alpha \leqslant \beta < \infty$, 因此，算子 $\mathcal{L}$ 是一致椭圆性. 也假设初函数 $g \in C_b^{\infty}(\mathbb{R}^d)$.

定理 7.22　在上述假设成立的情况下，对于任意 $T > 0$, Cauchy 问题 (7.3.4) 存在唯一解 $u(x,t) \in C^{2,1}(\mathbb{R}^d \times (0,T)) \bigcap C^{\infty}(\mathbb{R}^d \times (0,T))$. 此外，对于任意 $t > 0$ 如下估计成立：

$$\|u\|_{L^{\infty}(\mathbb{R}^d \times [0,t])} \leqslant \|g\|_{L^{\infty}(\mathbb{R}^d)} + \int_0^t \|f(\cdot, s)\|_{L^{\infty}(\mathbb{R}^d)} \mathrm{d}s. \tag{7.3.6}$$

注 7.23　定理 7.22 中的解 $u(x,t)$ 是一个古典解. 估计 (7.3.6) 是抛物型偏微分方程极大值原理的一个结果 (见推论 7.21). Cauchy 问题 (7.3.4) 的弱形式是通过乘以定义在 $\mathbb{R}^d$ 上的一个光滑、带有紧支集的函数，并运用分部积分得到 (见练习第 8 题). 在弱解背景下，当 $\mathcal{L}$ 采用散度形式时，Cauchy 问题 (7.3.4) 的解也可能得到能量估计

$$\|u\|_{L^{\infty}((0,T);L^2(\mathbb{R}^d))} + C_1 \|\nabla u\|_{L^2((0,T);L^2(\mathbb{R}^d))} \leqslant C_2(T) \left(\|f\|_{L^2((0,T);L^2(\mathbb{R}^d))} + \|g\|_{L^2(\mathbb{R}^d)} \right). \tag{7.3.7}$$

下一节中将讨论 $\mathbb{R}^d$ 上双曲方程的弱解. □

7.4　运输偏微分方程

本节探讨双线性运输偏微分方程中 Cauchy 问题的解的一些基本性质，这些方程是双曲型偏微分方程或对流方程. 研究无散度形式下的问题 —— 特别是要研究如下方程：

$$\frac{\partial u}{\partial t} + a(x) \cdot \nabla u = f(x), \quad (x,t) \in \mathbb{R}^d \times (0,T), \tag{7.4.1a}$$

$$u = g(x), \quad (x,t) \in \mathbb{R}^d \times \{0\}, \tag{7.4.1b}$$

其中 $a \in C_b^2(\mathbb{R}^d;\mathbb{R}^d)$, $f,g \in H^1(\mathbb{R}^d)$. 注意到当 $f \equiv 0$ 时, (7.4.1) 是如下微分方程的向后方程:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -a(x),$$

现将给 (7.4.1) 的解定义一个合理的概念, 并给出存在性和唯一性定理.

对于弱解的定义, 由于 (7.4.1) 这类方程总是涉及一个测试函数的增值和 $\mathbb{R}^d$ 上的积分, 因此, 有如下定义:

定义 7.24　如果

- $\dfrac{\partial u}{\partial t} \in L^2\left((0,T);L^2(\mathbb{R}^d)\right)$;
- 对于每一个 $\phi(x) \in H^1(\mathbb{R}^d)$, $\left(\dfrac{\partial u}{\partial t},\phi\right) + (a(x)\cdot\nabla u,\phi) = (u,f)$, 其中 $(\cdot,\cdot)$ 表示 $L^2(\mathbb{R}^d)$ 的内积;
- $u(x,0) = g$,

则称函数 $u \in L^2\left((0,T);H^1(\mathbb{R}^d)\right)$ 是 (7.4.1) 的一个弱解.

注 7.25　定义 7.24 与有界域上抛物型偏微分方程的弱解的定义是一致的, 即定义 7.17. 在研究抛物型的情况下, 那个定义也可以拓展到无界域, 正如这里对运输方程做的一样. □

注 7.26　Sobolev 空间的特性暗含了 $u \in C\left([0,T];L^2(\mathbb{R}^d)\right)$, 并且因此使得方程 $u(x,0) = g$ 有意义.

双线性运输方程 (7.4.1) 的弱解的存在性和唯一性成立.

定理 7.27　设 $a(x) \in C_b^2(\mathbb{R}^d;\mathbb{R}^d)$, $f,g \in H^1(\mathbb{R}^d)$, 则 (7.4.1) 存在一个弱解且唯一, 而且如下先验估计成立:

$$\|u\|_{L^\infty((0,T);L^2(\mathbb{R}^d))} \leqslant C(T,f,g). \tag{7.4.2}$$

在向量场 $a(x)$ 为无散度的情况下, 使得 $\nabla\cdot a(x) = 0$, 若 $f \equiv 0$, 以上估计变成

$$\|u\|_{L^\infty((0,T);L^2(\mathbb{R}^d))} \leqslant \|g\|_{L^2(\mathbb{R}^d)}. \tag{7.4.3}$$

证明　在此只给出证明的大致框架, 它基于黏性消除法: 其思想就是通过加上一个空间的二阶导数

$$\frac{\partial u^\varepsilon}{\partial t} + a(x)\cdot\nabla u^\varepsilon - \varepsilon\Delta u^\varepsilon = f(x), \quad (x,t) \in \mathbb{R}^d \times (0,T), \tag{7.4.4a}$$

$$u^\varepsilon = g^\varepsilon(x), \quad (x,t) \in \mathbb{R}^d \times \{0\} \tag{7.4.4b}$$

来 "规范" 初值问题 (7.4.1), 其中 $\varepsilon > 0$ 且 g^ε 是光滑的, 紧支撑近似趋于 $g(x)$[①]. 这个初值问题的解的存在性和唯一性, 遵循空间 $L^\infty\left((0,T);H^1(\mathbb{R}^d)\right)$ 上 Banach 不

① 这个近似是通过提取 g 的卷积得到的, 其中 g 具有紧支撑函数, 函数 $g^\varepsilon(x)$ 被称为 $g(x)$ 的近似.

动点定理的应用. 由抛物型正则性定理, 以及对 a, f 和 g 的假设暗含了

$$u^{\varepsilon} \in L^2\left(0,T;H^3\left(\mathbb{R}^d\right)\right),\quad \frac{\partial u^{\varepsilon}}{\partial t} \in L^2\left((0,T);H^1\left(\mathbb{R}^d\right)\right),$$

通过解 u^{ε} 合适的泛函乘以 (7.4.4), 进行分部积分, 并规范不等式导致以下先验估计:

$$\left\|u^{\varepsilon}\right\|_{L^{\infty}((0,T);H^1(\mathbb{R}^d))} + \left\|\frac{\partial u^{\varepsilon}}{\partial t}\right\|_{L^{\infty}((0,T);L^2(\mathbb{R}^d))} \leqslant C\left(g,f\right). \tag{7.4.5}$$

更重要的是 $C(g,f)$ 独立于 ε, 这个估计暗含了可以提取出一个弱收敛子序列, 仍由 u^{ε} 表示, 使得

$$u^{\varepsilon} \rightharpoonup u \quad 在\ L^2\left((0,T);H^1\left(\mathbb{R}^d\right)\right) 上为弱收敛,$$

$$\frac{\partial u^{\varepsilon}}{\partial t} \rightharpoonup u \quad 在\ L^2\left((0,T);L^2\left(\mathbb{R}^d\right)\right) 上为弱收敛.$$

通过 (7.4.4) 弱形式的极限 $\varepsilon \to 0$ 可以得到 (7.4.1) 中的弱形式, 也可以运用测试函数 $v \in C\left([0,T];H^1\left(\mathbb{R}^d\right)\right)$, $v(x,T)=0$ 说明 $u(x,0)=g(x)$, 在 $(0,T)$ 上对弱形式 (7.4.1) 积分, 并对 t 进行分部积分. 为了证明唯一性, 假设存在两个解 u_1 和 u_2 并形成它们的差分 $u=u_1-u_2$, 可以求解如下齐次化初值问题:

$$\frac{\partial u}{\partial t} + a(x)\cdot\nabla u = 0,\quad (x,t)\in\mathbb{R}^d\times(0,T), \tag{7.4.6}$$

$$u=0,\quad (x,t)\in\mathbb{R}^d\times\{0\}, \tag{7.4.7}$$

也导致

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2(\mathbb{R}^d)}^2 + \int_{\mathbb{R}^d}(a\cdot\nabla u)\,u\mathrm{d}x = 0.$$

接着设 $v \in C_0^{\infty}\left(\mathbb{R}^d\right)$, 计算

$$\begin{aligned}\int_{\mathbb{R}^d}(a(x)\cdot\nabla v)v\mathrm{d}x &= \frac{1}{2}\int_{\mathbb{R}^d}a(x)\cdot\nabla v^2\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}^d}(\nabla\cdot a)v^2\mathrm{d}x \\ &\leqslant \frac{1}{2}\|\nabla\cdot a\|_{L^{\infty}(\mathbb{R}^d)}\|v\|_{L^2(\mathbb{R}^d)}^2 \leqslant C\|v\|_{L^2(\mathbb{R}^d)}^2.\end{aligned} \tag{7.4.8}$$

一个近似参数变化可导出差分 $u(x,t)$ 满足如下估计:

$$\int_{\mathbb{R}^d}(a(x)\cdot\nabla u)\,u\mathrm{d}x \leqslant C\|u\|_{L^2(\mathbb{R}^d)}^2.$$

因此,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2(\mathbb{R}^d)}^2 \leqslant C\|u\|_{L^2(\mathbb{R}^d)}^2.$$

Gronwall 不等式给出 $u\equiv 0$, 从而得出其唯一性.

现在进行估计 (7.4.2) 和 (7.4.3) 的证明. 在 (7.4.1a) 上乘以 u, 在 $\mathbb{R}^d$ 上积分, 并用估计 (7.4.8) 和 Cauchy-Schwarz 不等式可得

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2(\mathbb{R}^d)}^2 &\leqslant \|f\|_{L^2(\mathbb{R}^d)}\|u\|_{L^2(\mathbb{R}^d)} + C\|u\|_{L^2(\mathbb{R}^d)}^2 \\ &\leqslant \|f\|_{L^2(\mathbb{R}^d)}^2 + C\|u\|_{L^2(\mathbb{R}^d)}^2 .\end{aligned}$$

由 Gronwall 引理可得

$$\|u\|_{L^2(\mathbb{R}^d)}^2 \leqslant \mathrm{e}^{Ct}\left(\|g\|_{L^2(\mathbb{R}^d)}^2 + \|f\|_{L^2(\mathbb{R}^d)}^2\right).$$

估计 (7.4.2) 可以在 $[0,T]$ 上通过对上述不等式两边取上确界得到.

为了证明估计 (7.4.3), 注意到 $a(x)$ 实际上是无散度的, 分部积分导致 (7.4.8), 这暗含了

$$\int_{\mathbb{R}^d}(a\cdot\nabla u)u\mathrm{d}x = 0.$$

设 $f\equiv 0$, 因此, 可得

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2(\mathbb{R}^d)}^2 \leqslant 0,$$

从以上可以看出估计成立.

注 7.28 设 u 是 (7.4.1) 的唯一弱解, 乘以一个测试函数 $\phi\in C_0^\infty\left(\mathbb{R}\times\mathbb{R}^d\right)$, 并在 $\mathrm{R}^+\times\mathrm{R}^d$ 上对 x 和 t 进行分部积分可得

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi}{\partial t}+\nabla\cdot(a\phi)\right)u\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}g(x)\phi(x,0)\,\mathrm{d}x = 0. \tag{7.4.9}$$

方程 (7.4.9) 对于第 21 章中运输方程的平均理论的证明是非常有用的. 有时确实要涉及 u 作为一个弱解满足 (7.4.9). □

7.5 半　　群

在前面章节中遇到的抛物型偏微分方程和双曲型偏微分方程可以看成是一个函数空间上的常微分方程 (这个观点依据定义 7.17 和定义 7.24), 采取形式

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{L}u + f, \tag{7.5.1}$$

其中, 将 $\mathcal{L}$ 看成是函数空间 X 的一个算子. 为了准确性, 假设 $\mathcal{L}:\mathcal{D}(\mathcal{L})\subset X\to X$. 在本章中, 将 $\mathcal{L}$ 看成是在函数 $u(x,t)$ 上每一点 (x,t) 都起作用的算子, 采用这个观点, 即 $\mathcal{L}$ 作用于一个函数空间就形成一个偏微分方程, 就像是函数空间中的常微分方程一样. 这就是探讨演化偏微分方程的半群方法.

正式地讲, 从所学知识知道 X 是有限维的情况下, 可以推导出 (7.5.1) 的解为

$$u(t)=\mathrm{e}^{\mathcal{L}t}u(0)+\int_0^t \mathrm{e}^{\mathcal{L}(t-s)}f(s)\,\mathrm{d}s. \tag{7.5.2}$$

这种表示被称为常数变异表示. 积分公式可以看成一个定义的起始点, 这个定义就是 (7.5.1) 的适定解. 要进行这个项目需要算子 $\mathrm{e}^{\mathcal{L}t}$ 的定义. 根据背景不同, 这个问题也有不同的解决方法, 很多运用 Fourier 变换、Fourier 级数及其推广.

为了说明这个观点, 考虑这种情况: X 是 Hilbert 空间, $\mathcal{L}$ 有一组特征函数, 它形成 X 的标准正交基, 可以表示为

$$-\mathcal{L}\phi_j=\lambda_j\phi_j,\quad \langle\phi_i,\phi_j\rangle=\delta_{ij},$$

其中, δ_{ij} 表示 Kronecker 三角张量, $\langle\cdot\rangle$ 为 H 上的内积, 可以通过

$$u=\sum_{j=1}^{\infty}u_j\phi_j,$$

$$\mathrm{e}^{\mathcal{L}t}u=\sum_{j=1}^{\infty}\mathrm{e}^{\lambda_j t}u_j\phi_j$$

来定义 $\mathrm{e}^{\mathcal{L}t}$. 这个级数的收敛性取决于 $\mathcal{L}$ 的谱性质以及 u 的正则性 (表现为 u_j 随着 j 的衰减). 一旦 $\mathrm{e}^{\mathcal{L}t}$ 被定义且 H 是 Hilbert 空间, 则通过

$$\left(\mathrm{e}^{\mathcal{L}t}\right)^*=\mathrm{e}^{\mathcal{L}^*t}$$

使得它的伴随有意义.

7.6 讨论和参考

本章中的内容是非常普通的, 可以在很多偏微分方程泛函分析的书中找到, 如文献 [53, 98, 121, 274]. 椭圆型 Dirichlet 问题的处理紧随 [98, 第 6 章]. 周期边界条件下对椭圆型偏微分方程的讨论是基于 [66, 第 4 章]. Dirichlet 问题的 Fredholm 理论在 [98, 6.2.3 小节] 得以推广.

数据在有足够正则性的情况下可以使得 Dirichlet 问题 (7.2.1) 得到一个古典解 (即函数 $u\in C^2(\Omega)\bigcap C(\overline{\Omega})$ 满足 (7.2.1)), 则弱解和古典解一致. 例如, 见文献 [121].

可以看到, 本章中对散度形式下的算子分析是基于适当函数空间上的能量法. 另一方面, 无散度形式下的偏微分方程基于极大值原理更合适. 椭圆型偏微分方

程极大值原理的研究见文献 [121], 而抛物型偏微分方程极大值原理的研究见文献 [264], 也可见文献 [98]. 当然, 只要系数是 C^1, 通过介绍涉及一阶导数的项, 就可以将散度形势下的偏微分方程改写为无散度形式下的偏微分方程, 反之亦然. 无散形式下的算子很自然地出现在扩散概率理论中, 如 Markov 过程中的生成元. 正如 6.5 节中提到的, 一个无散度形式二阶一致椭圆性算子中零空间 L^2 伴随矩阵的维数与 Markov 过程的遍历性理论有关, 见文献 [241, 246].

接下来转向抛物型偏微分方程, 定理 7.22 的证明可以在文献 [113] 中找到, 也可以见文献 [112, 304]. 具有时变系数的抛物型偏微分方程的类似理论也成立, 也可以说明 Cauchy 问题弱解的概念, 那么也可以运用 Banach 不动点定理证明解的存在性和唯一性, 见文献 [98, 7.3.2b]. 抛物型偏微分方程也可以用概率方法来研究. 确实, 标注它是一个偏微分方程的一个向后 Kolmogorov 方程后, 当 $f \equiv 0$ 时, 对系数进行合适的假设, (7.3.4) 的解可由一个概率表示, 更一般地, 当椭圆型偏微分方程含有一个与那个解成正比的源项时, Feynman-Kac 公式就提供了一个概率表示. 抛物型偏微分方程的解的存在性和唯一性的概率证明见 [156]. 无界系数的抛物型偏微分方程的研究见文献 [58, 196].

7.4 节大部分是源自 [98, 7.3.2 小节], 其中给出了定理 7.27 的证明. 黏性消除法的运用, 运输偏微分方程存在性和唯一性定理都可以通过第 4 章介绍的特征法来证明. 由特征法构造的解是古典解. 7.4 节中研究的弱解对于第 21 章中的平均化问题更为合适.

研究时变性偏微分方程的半群方法的综述见文献 [259]. 用 Hille-Yosida 定理证明的耗散抛物型偏微分方程的存在性和唯一性定理, 以及运用 Trotter-Kato 定理证明的极限定理是核心. 半群的运用是随机过程中极限定理这一严格理论的核心, 见文献 [94, 181～183].

7.7 练　习

1. 设 $\mathcal{A}$ 是散度形式的椭圆型算子 (7.2.3), 并设 $A(x) \in C^1(\Omega; \mathbb{R}^d)$, 将 $\mathcal{A}$ 转化为无散度形式 (7.2.4).

2. 以弱形式 (7.2.7) 开始, 证明注 7.4 中说明的结论.

3. 使用 Lax-Milgram 定理证明定理 7.8.

4. 使用定理 7.9 推导定理 7.8 的存在性和唯一性.

5. 说明并证明周期问题 (7.2.2) 的一个类似于注 7.4 中讨论的结论.

6. 在类似 (7.2.16) 的假设条件下, 证明 (7.2.14) 中定义的算子 $\mathcal{A}$ 的 Fredholm 性质只适用于 Dirichlet 边界条件.

7. 运用类似于推论 7.21 所用到的证明方法证明推论 7.16.

8. 考虑抛物型偏微分方程

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f(x), \quad (x,t) \in \Omega \times [0,T], \tag{7.7.1a}$$

$$u = 0, \quad (x,t) \in \partial\Omega \times (0,T], \tag{7.7.1b}$$

$$u = g, \quad (x,t) \in \bar{\Omega} \times \{0\}, \tag{7.7.1c}$$

其中

$$\mathcal{L} := \Delta \tag{7.7.2}$$

且 $\Omega \subset \mathbb{R}^d$ 是有界的. 制定一个弱解并证明方程有唯一稳态解 $\bar{u}(x)$. 通过能量估计证明当 $t \to \infty$ 时, $u(x,t) \to \bar{u}(x)$.

9. 仅当 $a = a(x)$ 和 $b = b(x)$ 的情况下, 运用特征法求解方程

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = b, \quad (x,t) \in \mathbb{R} \times \mathbb{R}^+,$$

$$u = 0, \quad (x,t) \in \mathbb{R} \times \{0\}.$$

10. 运用特征法求解 Burger's 方程

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad (x,t) \in \mathbb{R} \times (0,T),$$

$$u = g, \quad (x,t) \in \mathbb{R} \times \{0\}.$$

作为 g 的一个函数, 你认为 T 的值可以是多大值?

11. 考虑 Cauchy 问题 (7.3.4), 其中算子 $\mathcal{L}$ 是无散度形式. 假设存在唯一解在无穷远方向上衰减得足够快, 证明一个先验估计 (7.3.7), 并说明常数 $C_2(T)$ 是怎样取决于 T 的.

12. 考虑 Cauchy 问题

$$\frac{\partial u}{\partial t} = \nabla \cdot (A\nabla u), \quad (x,t) \in \mathbb{R}^d \times (0,T],$$

$$u = g, \quad (x,t) \in \mathbb{R}^d \times \{0\},$$

其中 $A = A(x)$ 满足标准假设, 说明其解在时间上是满足如下一致估计:

$$\|u\|_{L^\infty((0,T);L^2(\mathbb{R}^d))} + C_1 \|\nabla u\|_{L^2((0,T);L^2(\mathbb{R}^d))} \leqslant C_2 \|g\|_{L^2(\mathbb{R}^d)}.$$

13. 考虑椭圆型偏微分方程 (7.7.1), 其中 $\mathcal{L} = \Delta$, 定义域 $\Omega = [0,1]$ 且 $F \equiv 0$. 重新在 Hilbert 空间 $L^2(\Omega)$ 上形成常微分方程, 并且运用 Fourier 正弦级数找到一个合成半群的替代.

第二部分　扰 动 展 开

第8章　常微分方程的不变流形

8.1　引　　言

在动力系统中出现变量减少时, 也许这一种最简单的情形就是不变流形最吸引人的地方, 这些流形使变量的子集隶属于另外一个子集. 本章描述这样一种情形, 即在尺度分离系统中, 可以通过扰动展开构造有趣的不变流形. 在 8.2 节中引入常微分方程系统, 通过不变流形理论将其简化, 并在 8.3 节中给出简化方程. 在 8.4 节中推导简化方程, 并在 8.5 节中给出一些例子. 在 8.6 节中对本章结论进行延伸并给出参考书目, 最后还讨论了一些在第 9, 10 章中讲述的平均化问题.

8.2　完 全 方 程

考虑 (4.1.1) 形式的常微分方程系统, 记 $z=\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}$ 且满足如下方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f\left(x,y\right), \tag{8.2.1a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{1}{\varepsilon}g\left(x,y\right) \tag{8.2.1b}$$

且 $\varepsilon\ll 1$, 其中 $x\in\mathcal{X}$ 以及 $y\in\mathcal{Y}$ 采用第 4 章中的记号.

当将 x 视为固定参数且 $\varepsilon=1$ 时, 设 $\varphi_x^t\left(y\right)$ 为快速动力系统中的解算子. 确切地说, 对任意 $\xi\in\mathcal{X}$, 设

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_\xi^t\left(y\right)=g\left(\xi,\varphi_\xi^t\left(y\right)\right),\quad \varphi_\xi^0\left(y\right)=y. \tag{8.2.2}$$

假定

$$\lim_{t\to\infty}\varphi_\xi^t\left(y\right)=\eta\left(\xi\right) \tag{8.2.3}$$

存在, 其独立于 y 且在 ξ 上是一致收敛的. 大致来说, (8.2.1) 的解 $y\left(t\right)$ 由 $y\left(t\right)\approx\varphi_{x(0)}^{t/\varepsilon}\left(y\left(0\right)\right)$ 给出, 其中时间 t 与 1 比较是一个无穷小量 (即 $t=o(1)$), 这就使得 $x\left(t\right)$ 还没有多大变化. 如果认为与 ε 相比, 短时间尺度是大的, 使得 y 接近其均衡点 (如若 $t=O\left(\varepsilon^{1/2}\right)$), 就可得 $y\left(t\right)\approx\eta\left(x\left(0\right)\right)$. 通过这种原理使得 y 隶属于 x, 并且需要寻找这种原理使得这个启发式的想法更为准确.

注意到 (8.2.1) 的生成子 $\mathcal{L}$ 有如下形式:

$$\mathcal{L} = \frac{1}{\varepsilon}\mathcal{L}_0 + \mathcal{L}_1, \tag{8.2.4}$$

其中

$$\mathcal{L}_0 = g(x,y)\cdot\nabla_y, \quad \mathcal{L}_1 = f(x,y)\cdot\nabla_x.$$

特别地, 对每个固定 x, 在 $\mathcal{Y}$ 上, $\mathcal{L}_0$ 是一个过程的生成子.

对 $v(y,t)$ 考虑下面的偏微分方程, 其中 x 视为一个固定的参数:

$$\frac{\partial v}{\partial t} = \mathcal{L}_0 v, \quad v(y,0) = \phi(y). \tag{8.2.5}$$

定理 4.6 表明

$$v(y,t) = \phi\left(\varphi_x^t(y)\right).$$

从而由 (8.2.3),

$$v(y,t) \to \phi(\eta(x)), \quad t\to\infty. \tag{8.2.6}$$

这与遍历性有关, 因为方程 (8.2.3) 表明随着 $t\to\infty$, 函数 $v(y,t)$ 不独立于初值并接近于 y. 与第 4 章中对遍历性的讨论比较一下, 尤其是定理 4.12 和定理 4.13.

回想不变集合的定义 4.3, 如果集合是流形的, 则称之为不变流形. 本章中使用方程 (8.2.1) 中的尺度分离形式构造一个近似不变流形. 事实上, 流形具有图形结构, 可以表示为一个坐标 y 与坐标 x 的函数. 在描述常微分方程动力系统接近平衡时, 将不变流形表示为图形是很重要的, 不变流形还提出了稳定、不稳定以及中心流形的概念.

8.3 简化方程

现在说明一个由下节形式扰动的观点得出的近似结论. 由下式定义向量场 $F_0(x)$:

$$F_0(x) = f(x,\eta(x)). \tag{8.3.1}$$

结论 8.1　对于 $\varepsilon \ll 1$ 且时间 t 具有同阶 $O(1)$, (8.2.1) 的解 $x(t)$ 近似于下面方程的解 $X(t)$:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F_0(X), \tag{8.3.2}$$

其中 $F_0(x)$ 由 (8.3.1) 给出.

下面对结论 8.1 的推导是假设开始时 $y(0)$ 是接近于 $\eta(x(0))$ 的. 如果这个假设不成立, 则需要进一步研究何为初始层或者边界层, 关于这一点的讨论见 8.6 节.

结论 8.1 给出了 ε 的首阶近似, 如果保持 ε 的高一阶, 就能得到更为精确的近似值

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F_0(X) + \varepsilon F_1(X), \tag{8.3.3}$$

此处:

$$F_1(X) = \nabla_y f(x,\eta(x))\left(\nabla_y g(x,\eta(x))\right)^{-1} \nabla_x \eta(x) f(x,\eta(x)).$$

这个近似值要求 $\nabla_y g(x,\eta(x))$ 是可逆的.

8.4 推　导

求简化方程的方法就是寻找一个系统的近似不变流形, 而且假定流形可由关于 x 的图形表示, 即 $y = \Psi(x)$. 在动力系统下, 对任何 $y = \Psi(x)$, 如果

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \nabla\Psi(x(t))\frac{\mathrm{d}x}{\mathrm{d}t},$$

则称由此图形决定的集合是不变的 (见定义 4.3). 这暗含 Ψ 一定是如下非线性偏微分方程的解:

$$\frac{1}{\varepsilon} g(x, \Psi(x)) = \nabla\Psi(x) f(x, \Psi(x)).$$

需要寻找上述方程幂级数形式的解:

$$\Psi(x) = \Psi_0(x) + \varepsilon\Psi_1(x) + O\left(\varepsilon^2\right).$$

这就是扰动展开的第一个例子.

代入上述方程, 比较方程两边 ε 的各阶次数的系数得到分层链

$$O\left(\frac{1}{\varepsilon}\right) : g(x, \Psi_0(x)) = 0;$$
$$O(1) : \nabla_y g(x, \Psi_0(x))\Psi_1(x) = \nabla\Psi_0(x) f(x, \Psi_0(x)).$$

注意到方程 (8.2.2) 和 (8.2.3) 都暗含了对于所有的 ξ, $g(\xi,\eta(\xi)) = 0$, 因而令 $\Psi_0(x) = \eta(x)$, 可以满足 $O(1/\varepsilon)$ 方程, 也就得到 (8.3.2) 的近似值. 由于假设 (8.2.3) 中收敛速率是一致的, 通常假定 $y = \eta(\xi)$ 是 (8.2.2) 的一个双曲平衡点[①], 使得 $\nabla_y g(x,\eta(x))$ 是可逆的. 在 $O(1)$ 方程中设 $\Psi_0(x) = \eta(x)$, 求逆得

$$\Psi_1(x) = \nabla_y g(x,\eta(x))^{-1} \nabla\eta(x) f(x,\eta(x)).$$

① 双曲平衡点就是在平衡点处, 向量场的线性化在假象轴上不包含谱.

因此,

$$
\begin{aligned}
f(x, \Psi(x)) &= f\left(x, \Psi_0(x) + \varepsilon\, \Psi_1(x) + O\left(\varepsilon^2\right)\right) \\
&= f(x, \Psi_0(x)) + \varepsilon\, \nabla_y f(x, \Psi_0(x))\, \Psi_1(x) + O\left(\varepsilon^2\right) \\
&= f(x, \eta(x)) + \varepsilon\, \nabla_y f(x, \eta(x))\, \Psi_1(x) + O\left(\varepsilon^2\right),
\end{aligned}
$$

并且得到 (8.3.3) 的精确值.

8.5 应　　用

8.5.1 线性快速动力学

在许多应用中出现了一种结构, 在该结构中, 由 $\varphi_\xi^t(\cdot)$ 给出的 x 是线性的. 作为简单的例子, 考虑如下方程:

$$
\begin{cases}
\dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x, y), \\
\dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{y}{\varepsilon} + \dfrac{\tilde{g}(x)}{\varepsilon},
\end{cases} \tag{8.5.1}
$$

其中 $d = 2$ 且 $\mathcal{X} = \mathcal{Y} = \mathbb{R}, \mathcal{Z} = \mathbb{R}^2$. 显然有

$$
\varphi_\xi^t(y) = \mathrm{e}^{-t} y + \int_0^t \mathrm{e}^{s-t} \tilde{g}(\xi)\, \mathrm{d}s = \mathrm{e}^{-t} y + \left(1 - \mathrm{e}^{-t}\right) \tilde{g}(\xi),
$$

因而对 $\eta(\cdot) = \tilde{g}(\cdot)$, (8.2.3) 满足.

因此, 由结论 8.1 给出的简化方程为

$$
\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, \tilde{g}(X)).
$$

由 $\nabla_y g(x, y) = -1$ 得到 (8.3.3) 更精确的近似值为

$$
\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, \tilde{g}(X)) \left(1 - \varepsilon \frac{\mathrm{d}f}{\mathrm{d}y}(X, \tilde{g}(X)) \frac{\mathrm{d}\tilde{g}}{\mathrm{d}x}(X)\right).
$$

8.5.2 长时间动力学

简化动力系统的结论关注了 $x(t)$ 关于 ε 在 $O(1)$ 时间间隔上的近似值. 然而, 在许多情况下, 可以很自然地将结论延伸到无限时间区域上. 下面的例子说明了这个思想.

考虑方程

$$\begin{cases}\dfrac{\mathrm{d}x_1}{\mathrm{d}t}=-x_2-x_3, & (8.5.2a)\\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t}=x_1+\dfrac{1}{5}x_2, & (8.5.2b)\\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t}=\dfrac{1}{5}+y-5x_3, & (8.5.2c)\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=-\dfrac{y}{\varepsilon}+\dfrac{x_1x_3}{\varepsilon}, & (8.5.2d)\end{cases}$$

使得 $\mathcal{X}=\mathbb{R}^3$ 且 $\mathcal{Y}=\mathbb{R}$. 结论 8.1 表示用 X 求解下面的 Rössler 系统可以很好地得到 x 的近似值:

$$\begin{cases}\dfrac{\mathrm{d}X_1}{\mathrm{d}t}=-X_2-X_3, & (8.5.3a)\\ \dfrac{\mathrm{d}X_2}{\mathrm{d}t}=X_1+\dfrac{1}{5}X_2, & (8.5.3b)\\ \dfrac{\mathrm{d}X_3}{\mathrm{d}t}=\dfrac{1}{5}+X_3(X_1-5). & (8.5.3c)\end{cases}$$

由于 Rössler 方程是混沌的, 因此, 在长时间间隔上比较轨迹是不自然的. 吸引子是一个更加有用的话题, 比较用数值法生成的两个系统的吸引子, 如图 8.1 所示. 左图为 $\varepsilon=10^{-2}$ 时投影到 x 轴上的方程 (8.5.2) 的吸引子, 右表为 Rössler 方程本身的吸引子. 两幅图的一致性充分表明简化动力系统的确能捕捉到在长时间间隔上的特性.

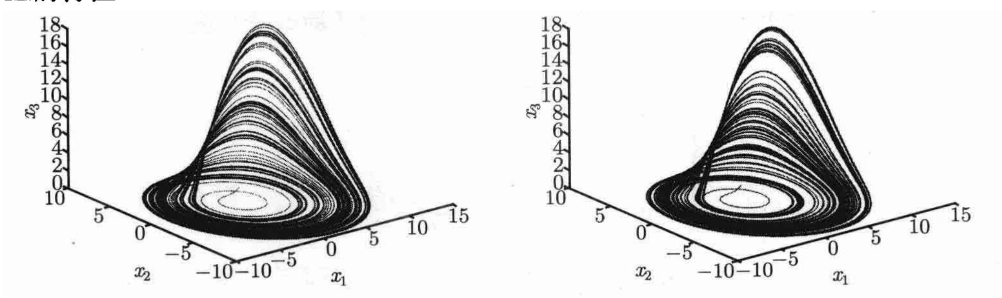


图 8.1　分别投影到 (x_1,x_2) 和 (X_1,X_2) 平面上时方程 (8.5.2)($\varepsilon=0.01$ 左图) 与方程 (8.5.3) 吸引集合的比较

8.5.3　中心流形

中心流形是一个包含平衡点的不变流形, 该平衡点的线性化具有中性方向 (对应于实部为 0 的特征值的子空间). 考虑方程

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=\lambda x+\displaystyle\sum_{i=0}^{2}a_ix^iy^{2-i},\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=x-y+\displaystyle\sum_{i=0}^{2}b_ix^iy^{2-i},\end{cases}$$

其中 $\lambda \in \mathbb{R}$, a_i 和 b_i 都为实数, 而且对每个 $t, x(t) \in \mathbb{R}$ 以及 $y(t) \in \mathbb{R}$, 在起点将其线性化, 方程变为

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \lambda x, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = x - y. \end{cases}$$

如果 $z = (x, y)^{\mathrm{T}}$, 则

$$\frac{\mathrm{d}z}{\mathrm{d}t} = Lz,$$

其中

$$L = \begin{pmatrix} \lambda & 0 \\ 1 & -1 \end{pmatrix}.$$

L 的特征值为 λ 和 -1. 因为 λ 穿过 0, 因而初始点的线性稳定性由稳定变为不稳定. 因此, 在 $\lambda = 0$ 附近研究方程是很有趣的. 尤其希望在 $\lambda = 0$ 处找的一个中心流形: 一个与 L 的特征值为 0 时所对应的特征空间相切的不变流形.

为了构造这样的流形, 改变尺度可得方程如下: 设

$$x \to \varepsilon x, \quad y \to \varepsilon y, \quad \lambda \to \varepsilon \lambda, \quad t \to \varepsilon^{-1} t.$$

这相当于找一个小振幅解, 接近初始处的固定点, 参数值接近分叉值. 这样的解进化很慢, 为了捕捉不平凡的动力系统, 因此, 时间需要重新改变尺度. 方程变为

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \lambda x + \displaystyle\sum_{i=0}^{2} a_i x^i y^{2-i}, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\varepsilon}(x - y) + \displaystyle\sum_{i=0}^{2} b_i x^i y^{2-i}. \end{cases}$$

扰动展开给出不变流形 $y = x$, 并且在不变流形上得到下面动力系统的方程:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \lambda X + AX^2,$$

其中 $A = \displaystyle\sum_{i=0}^{2} a_i$. 当 $\lambda = 0$ 时给出自身的中心流形, 当 $\lambda < 0$ 时给出稳定流形.

8.6 讨论和参考

不变流形这个话题已经有很长的历史了, 并且它也是全书的主题. 详细地介绍不变流形是不可能的, 仅在文献中列出一些要点. 从全书的角度来看, 最初讨论这个话题是想提供一个关于平均算法 (在下面两章中介绍) 的例子, 进而可以在不掌

握遍历性理论以及随机过程理论的情况下就能引入这个例子, 因而为那些具有动力系统背景的读者提供一个合适的方法去理解本书中有关不变流形的话题; 反过来, 也给出了平均化与动力系统之间的具体关系, 在第 10 章将进一步讨论这个话题. 注意到从更高的层次来讲, 本章中用到的扰动展开与第 II 部分中用到的类似. 然而, 在一个重要的方面, 这个观点是不一样的, 即所有剩余章节都涉及线性问题近似值的扰动展开 (通过使用向后方程), 并依赖于反复使用 Fredholm 性质. 在本章中, 因为曲线 Ψ 是非线性的, 并且不能使用 Fredholm 性质, 所以使用扰动展开的方法会有些不同.

不变流形的一般理论见文献 [131, 331], 这些书中都特别强调了对方程 (4.1.1) 中不变集合的非稳定、稳定以及中心流形的构造. 特别地, 对于最简不变集合, 即一个平衡点, 可能会改变坐标系, 使得原点 0 是一个平衡点, 这样 (4.1.1) 形式如下:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = Lz + h_1(z), \quad z(0) = z_0,$$

其中与 $z \to 0$ 相比, $h_1(z)$ 是很小的. 在双曲平衡点上, L 的不变子空间分离为稳定空间和不稳定空间. 如果设 P 为稳定空间上的正交投影, 并且 $Q = I - P$ 为不稳定空间上的正交投影, 则记 $x = Pz, y = Qz$, 得到方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = L_1 x + f_1(x, y),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = L_2 y + g_1(x, y).$$

稳定流形 (在原点附近) 可由曲线 $y = \Theta(x)$ 表示, 同样, 不稳定流形可由曲线 $x = \Phi(y)$ 表示. 但是只有当 PZ 在 L 空间上含有中性方向时, 中心流形才类似于稳定流形. 在文献 [57] 中对中心流形作了特别的讨论. 由于 L 的中性谱中包含一对复杂的共轭特征值, 从而得到 Hopf 分叉定理, 这一特殊例子见文献 [212].

这些关于固定点附近行为的特殊不变流形定理表明曲线在其中起到很重要的作用, 这些曲线将变量的一个子集与构造不变流形时的另外一个子集联系起来. 这样的曲线图在构造所谓的 (8.2.1) 慢流形时起到很关键的作用. 早期对常微分方程近似值的研究包括了对 Levinson 和 Tikhonov 的研究, 通过微分代数方程这些常微分方程具有吸引慢流形 (见文献 [239, 317]). 在 8.3 节中提到近似结果的最简形式要求快速变量 y 在开始时就接近于不变流形. 然而, 即使不是这样, 也可以引入一个初始层 (有时称之为边界层) 来延伸近似值, 并通过相应的渐近展开式进行研究, 见文献 [71, 272, 324].

构造不变流形时, 通过使用 (8.2.2) 的渐近稳定固定点, 使得 y 显式地隶属于 x. 一般来说, 在构造局部稳定、非稳定、中心流形 (如文献 [57, 331]), 慢流形[178] 以

及惯性流形[69] 时使用与非线性项大小密切相关的谱隙. 特别地, 慢性流形的构造表明了如何在耗散偏微分方程中将不变流形理论推广到无限维空间上的思想.

关于动力系统的数值方法, 尤其是计算不变流形的数值方法的参考文献见 [305]; 也可以为随机 (偏) 微分方程构造不变流形, 见文献 [38, 39, 41, 77, 328, 329] 及其参考书.

8.7 练　习

1. 考虑方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \lambda x + a_0 x^3 + a_1 xy, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -y + \displaystyle\sum_{i=0}^{2} b_i x^i y^{2-i}, \end{cases}$$

其中 $\lambda \in \mathbb{R}$ 且 a_i 和 b_i 也是实数. 对每个固定的 t 有 $x(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$. 显示把尺度

$$x \to \varepsilon x, \quad y \to \varepsilon^2 y, \quad \lambda \to \varepsilon^2 \lambda, \quad t \to \varepsilon^{-2} t$$

应用到这个系统时所采用的扰动技巧. 推导中心流形具有如下形式:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \lambda X + AX^3,$$

其中 $A = a_0 + a_1 b_2$.

2. 假设 $\varepsilon > 0, A \in \mathbb{R}^{l\times l}$, 并且 $B \in \mathbb{R}^{(d-l)\times(d-l)}$. 对 $\varepsilon \ll 1$ 且 $x \in \mathbb{R}^l, y \in \mathbb{R}^{d-l}$, 考虑方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = Ax + \varepsilon f_0(x, y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{1}{\varepsilon} By + g_0(x, y), \end{cases}$$

假设 B 是对称正定的. 找出不变流形展开式中的前三项, 该流形表示 y 关于 x 的一个图形.

3. 假设 $\varepsilon > 0$ 且 $B \in \mathbb{R}^{(d-l)\times(d-l)}$. 对 $\varepsilon \ll 1$ 并且 $x \in \mathbb{R}^l, y \in \mathbb{R}^{d-l}$, 考虑方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x, y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{1}{\varepsilon}(By - \tilde{g}(x)). \end{cases}$$

a) 假定 B 是对称正定的, 找出不变流形展开式中的首项, 该流形表示了 y 关于 x 的一个图形;

b) 考虑 $d - l = 2, \tilde{g} \equiv 0$ 的情形, 并且

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

当 $\varepsilon \to 0$ 时, 解会怎样?

第 9 章　Markov 链的平均化

9.1　引　　言

也许揭露随机动力系统问题的消元法最简单的背景就是 Markov 链. 在这种背景下, 非常自然地需要研究这样一种情况: 变量的子集与余量相比进化得非常迅速, 并且这种变量子集的影响可以用它们的平均作用来代替. 在 9.2 节中描述了非平均化 Markov 链; 9.3 节给出了平均化方程; 平均化方程的导出见 9.4 节; 9.5 节中给出了一个例子; 9.6 节讨论了本章结论的推广并作了一些文献标注.

9.2　完 全 方 程

根据第 5 章中的结构, 考虑逆向方程

$$\frac{\mathrm{d}v}{\mathrm{d}t} = Qv. \tag{9.2.1}$$

回顾这个方程, 其中 $v(0) = \phi$, 则有如下性质:

$$v_i(t) = \mathbb{E}\left(\phi_{z(t)} \,|z(0) = i\right),$$

其中 $\mathbb{E}$ 表示相对于 Markov 转移概率的期望. 假设生成元 Q[①]采用形式

$$Q = \frac{1}{\varepsilon}Q_0 + Q_1, \tag{9.2.2}$$

其中 $0 < \varepsilon \ll 1$. 研究如下情况: 状态空间包含两个变量 x 和 y, Q 中的首项, 即 Q_0, 相当于 x 是固定时, y 方向上的快速遍历动力系统. 对 y 的平均值则给出 x 的有效动力性.

确切的情况如下: 状态空间是 $\mathcal{I} := \mathcal{I}_x \times \mathcal{I}_y$, 其中 $\mathcal{I}_x, \mathcal{I}_y \subseteq \{1, 2, \cdots\}$. 设 $q((i,k),(j,l))$ 表示生成元的一个元素, 这个生成元与从 $(i,k) \in \mathcal{I}_x \times \mathcal{I}_y$ 到 $(j,l) \in \mathcal{I}_x \times \mathcal{I}_y$[②]的转移有关. 考虑 $\mathcal{I}_y$ 上的一个 Markov 链族, 由 $i \in \mathcal{I}_x$ 引出. 把生成元写成 $A_0(i)$ 且有分量 $a_0(k,l;i)$, 这个分量表示对固定的 $i \in \mathcal{I}_x$, 从 $k \in \mathcal{I}_y$ 到 $l \in \mathcal{I}_y$

① 本章中用 Q 表示生成元, 而不是用 L, 这是因为用索引 l 表示状态空间, 从而避免与生成元的参数混淆.

② 本章和第 16 章中, 不用 suffices 表示状态空间上的独立性, 这是因为双索引使得它成为一个杂波符号, 因此, 使用 $q((i,k),(j,l))$, 而不是 $q_{(i,k),(j,l)}$.

的转移. 假定对于每个 $i \in \mathcal{I}_x$, $A_0(i)$ 在 $\mathcal{I}_y$ 上生成遍历性 Markov 链. 因此, 对于固定的 i,$A_0(i)$ 有一个一维的零空间, 并且[①]

$$\begin{aligned}\sum_l a_0(k,l;i) &= 0, \quad (i,k) \in \mathcal{I}_x \times \mathcal{I}_y,\\ \sum_k \rho^\infty(k;i)\, a_0(k,l;i) &= 0, \quad (i,l) \in \mathcal{I}_x \times \mathcal{I}_y.\end{aligned} \tag{9.2.3}$$

这里采用的是 (5.6.2) 的指标形式, 其中 $A_0(i)$ 代替了 L. 不失一般性, 选择规范化

$$\sum_k \rho^\infty(k;i) = 1, \quad \forall i \in \mathcal{I}_x,$$

从而 $\rho^\infty(i) = \{\rho^\infty(k;i)\}_{k\in\mathcal{I}_y}$ $(i \in \mathcal{I}_x)$ 是 $\mathcal{I}_y$ 上一个 Markov 链的不变分布.

类似于前面对 $\mathcal{I}_x$ 上 Markov 链生成子的介绍, 可知当 $k \in \mathcal{I}_y$ 时的参数化形式. 用 $A_1(k)$ 表示对于每个 $k \in \mathcal{I}_y$ 具有分量 $a_1(i,j;k)$ 的生成子, 其中 $a_1(i,j;k)$ 表示从 $i \in \mathcal{I}_x$ 到 $j \in \mathcal{I}_x$ 的转移. 通过

$$\begin{aligned}q_0((i,k),(j,l)) &= a_0(k,l;i)\,\delta_{ij},\\ q_1((i,k),(j,l)) &= a_1(i,j;k)\,\delta_{kl},\end{aligned} \tag{9.2.4}$$

对于符号 A_0, A_1, 引进 Q_0, Q_1 来表示 $\mathcal{I}_x \times \mathcal{I}_y$ 上 Markov 链的生成子, 其中 δ_{ij} 是常用 Kronecker 符号. 在 Q_0(或 Q_1) 的构造中, Kronecker 符号表示这样一个事实, 即在 $\mathcal{I}_x$(或 $\mathcal{I}_y$) 中没有转移发生.

为了要证实 Q_0, Q_1 是如定义中所述的生成子, 注意到由于 A_0 和 A_1 都是生成子, 则非对角元 $(i,k) \neq (j,l)$ 是非负的, 同样由 (9.2.3) 可得

$$\sum_{j,l} q_0((i,k),(j,l)) = \sum_{j,l} a_0(k,l;i)\,\delta_{ij} = \sum_l a_0(k,l;i) = 0.$$

类似地, 由于对每个固定的 k, $A_1(k)$ 是一个生成子, 并且运用事实

$$\sum_j a_1(i,j;k) = 0, \quad \forall (i,k) \in \mathcal{I}_x \times \mathcal{I}_y$$

可计算证明

$$\sum_{j,l} q_1((i,k),(j,l)) = 0.$$

这样 Q_0, Q_1 也是 Markov 链的生成子. 最后注意到生成子乘以正常数也是生成子, 因此, 生成子的任一线性组合也是一个生成子, 这样对于任意 $\varepsilon > 0$, (9.2.2) 定义了一个生成子.

① 本章中的总结总是关于 $\mathcal{I}_x$ 或 $\mathcal{I}_y$ 中的索引, 它应该是按照背景求和.

9.3 简化方程

通过

$$\overline{q}_1(i,j)=\sum_k \rho^{\infty}(k;i)\,a_1(i,j;k) \tag{9.3.1}$$

定义 $\mathcal{I}_x$ 上 Markov 链的一个生成子 $\overline{Q}_1$, 由于当 $i \neq j$ 时可知 $\rho^{\infty}(k;i) \geqslant 0$ 和 $a_1(i,j;k) \geqslant 0$, 因此有 $\overline{q}(i,j) \geqslant 0 (i \neq j)$, 而且

$$\sum_j \overline{q}_1(i,j)=\sum_k \rho^{\infty}(k;i)\left(\sum_j a_1(i,j;k)\right)=0.$$

这样 $\overline{Q}_1$ 是 Markov 链的生成子.

结论 9.1　考虑假设 (9.2.2) 情况下的方程 (9.2.1), 则对于 $\varepsilon \ll 1$ 和乘以 t 达到 $O(1)$ 时, 可以通过具有生成子 $\overline{Q}_1$ 的 Markov 链 X 近似于 $x \in \mathcal{I}_x$ 的有限维分布.

强调的是 x 并不是它自身的 Markovian, 仅仅对 (x,y) 是. 正如前面讨论的, $\overline{Q}_1$ 只是 $\mathcal{I}_x$ 上 Markov 链的一个生成子, 并且通过平均求值消除了 $\mathcal{I}_y$ 上的动力, 则近似变量 X 是 Markovian 且由如下向后方程决定:

$$\frac{\mathrm{d}v_0}{\mathrm{d}t}=\overline{Q}_1 v_0. \tag{9.3.2}$$

现在通过扰动展开的方式为这种变量的消除来提供正当的理由.

9.4 推　导

所用方法是为了说明当 $x \in \mathcal{I}_x$ 时的向后方程 (9.3.2) 可以近似于 $(x,y) \in \mathcal{I}_x \times \mathcal{I}_y$ 上的满 Markov 链的向后方程. 考虑 (9.2.2) 情况下的 (9.2.1), 可得如下向后方程:

$$\frac{\mathrm{d}v}{\mathrm{d}t}=\left(\frac{1}{\varepsilon}Q_0+Q_1\right)v.$$

不像先前的章节那样, 一个非线性偏微分方程近似含有一个很小的参数 ε, 这里的问题是线性的. 在接下来的 5 章中, 所有的扰动展开都是针对于类似的线性方程, 其典型推导如下.

寻求 $v=v(i,k,t)$ 的如下多尺度展开形式:

$$v=v_0+\varepsilon v_1+O\left(\varepsilon^2\right), \tag{9.4.1}$$

代替并使 ε 各阶项的系数为零可得

$$O\left(\frac{1}{\varepsilon}\right) Q_0 v_0 = 0, \tag{9.4.2a}$$

$$O(1)\, Q_0 v_1 = -Q_1 v_0 + \frac{\mathrm{d}v_0}{\mathrm{d}t}. \tag{9.4.2b}$$

由 (9.2.3) 可从 (9.4.2a) 中推导得知 v_0 独立于 $k \in \mathcal{I}_y$. 使用记号

$$v_0(i,k,t) = v_0(i,t)\, 1(k), \tag{9.4.3}$$

其中对于所有的 $k \in \mathcal{I}_y$ 都有 $1(k) = 1$. 算子 Q_0 是奇异的, 因此, 方程 (9.4.2b) 有一个解, Fredholm 性质暗含了可解条件

$$-Q_1 v_0 + \frac{\mathrm{d}v_0}{\mathrm{d}t} \perp \mathrm{Null}\left\{Q_0^{\mathrm{T}}\right\}. \tag{9.4.4}$$

从 (9.2.3) 可以推导出, 对于 $\mathcal{I}_x$ 上的任一向量 $c = \{c(i)\}$, 零空间 Q_0^{T} 可由

$$\sum_{k,i} \rho^{\infty}(k;i)\, c(i)\, q_0((i,k)(j,l)) = 0 \tag{9.4.5}$$

来表征. 由 (9.4.3) 可得

$$\begin{aligned}
\frac{\mathrm{d}v_0}{\mathrm{d}t} - Q_1 v_0 &= \frac{\mathrm{d}v_0}{\mathrm{d}t}(i,t)\, 1(k) - \sum_{j,l} a_1(i,j;k)\, \delta_{kl} v_0(j,t)\, 1(l) \\
&= \left(\frac{\mathrm{d}v_0}{\mathrm{d}t}(i,t) - \sum_j a_1(i,j;k)\, v_0(j,t)\right) 1(k).
\end{aligned}$$

通过 (9.4.5), 运用附加可解条件 (9.4.4) 可得

$$\sum_{k,i} \rho^{\infty}(k;i)\, c(i) \left(\frac{\mathrm{d}v_0}{\mathrm{d}t}(i,t) - \sum_j a_1(i,j;k)\, v_0(j,t)\right) = 0,$$

此式暗含了

$$\sum_i c(i) \left(\frac{\mathrm{d}v_0}{\mathrm{d}t}(i,t) - \sum_j \overline{Q}_1(i,j)\, v_0(j,t)\right) = 0.$$

由于 c 是 $\mathcal{I}_x$ 上的任意向量, 可以推导出 i 求和中的每个分量都是零. 这就导出了 (9.3.2).

9.5 应　用

考虑一个简单例子, 其中 $\mathcal{I}_x = \mathcal{I}_y = \{1,2\}$, 则在 $\mathcal{I}\mathcal{L} = \mathcal{I}_x \times \mathcal{I}_y$ 上有一个四态 Markov 链. 设 $\mathcal{I}_x$ 和 $\mathcal{I}_y$ 上的 Markov 链的生成子分别是由

$$A_0(i) = \begin{pmatrix} -\theta_i & \theta_i \\ \phi_i & -\phi_i \end{pmatrix}$$

和

$$A_1(k) = \begin{pmatrix} -\alpha_k & \alpha_k \\ \beta_k & -\beta_k \end{pmatrix}$$

给出的. 这些 Markov 链 $i \in \mathcal{I}_x$(或 $k \in \mathcal{I}_y$) 中的第一个 (或第二个) 是固定参数, 参数 $\theta_i, \phi_i, \alpha_k$ 和 β_k 都是非负的.

如果对 Markov 链的四态排序为 (1,1),(1,2),(2,1) 和 (2,2), 则生成子 Q_0 和 Q_1 分别为

$$Q_0 = \begin{pmatrix} -\theta_1 & \theta_1 & 0 & 0 \\ \phi_1 & -\phi_1 & 0 & 0 \\ 0 & 0 & -\theta_2 & \theta_2 \\ 0 & 0 & \phi_2 & -\phi_2 \end{pmatrix} \tag{9.5.1}$$

和

$$Q_1 = \begin{pmatrix} -\alpha_1 & 0 & \alpha_1 & 0 \\ 0 & -\alpha_2 & 0 & \alpha_2 \\ \beta_1 & 0 & -\beta_1 & 0 \\ 0 & \beta_2 & 0 & -\beta_2 \end{pmatrix}. \tag{9.5.2}$$

Q_0 和 Q_1 的任一线性组合沿着反对角线都是零, 因此, 同样对 Q 也成立. 由构造反应了这样一个事实: $\mathcal{I}_x$ 和 $\mathcal{I}_y$ 上的转移不会同时发生.

具有生成子 $A_0(i)$ 的 Markov 链的不变密度是 $\rho^\infty(i) = (\lambda_i, 1-\lambda_i)^{\mathrm{T}}$, 其中 $\lambda_i = \phi_i/(\theta_i+\phi_i)$. 回顾 $\mathcal{I}_x$ 上的平均化 Markov 链有生成子 $\overline{Q}_1$ 且它的分量可表示为

$$q_1(i,j) = \sum_k \rho^\infty(k;i)\, a_1(i,j;k) = \lambda_i a_1(i,j;1) + (1-\lambda_i)\, a_1(i,j;2),$$

则

$$\overline{Q}_1 = \begin{pmatrix} -\lambda_1\alpha_1 - (1-\lambda_1)\alpha_2 & \lambda_1\alpha_1 + (1-\lambda_1)\alpha_2 \\ \lambda_2\beta_1 + (1-\lambda_2)\beta_2 & -\lambda_2\beta_1 - (1-\lambda_2)\beta_2 \end{pmatrix}. \tag{9.5.3}$$

9.6 讨论和参考

对 Markov 链多尺度问题研究的两本最新专著见文献 [335, 336]; 随机动力学中对求平均值和维数降低的广泛讨论见文献 [291]; 对偏微分方程的 Markov 链逼近, 尤其是在大的极限偏差下的研究见文献 [111]; 对多尺度 Markov 链的计算方法的讨论见文献 [85, 86]; Markov 链驱动下常微分方程扩散极限的研究见文献 [245]; 与扩散极限理论相关的证明见文献 [96]; Markov 链和中心流形之间的关系见文献 [262].

本章中提出了 Markov 链的平均化. Markov 链的均匀化 (中心极限定理) 结论见文献 [184].

在推导近似方程中暗含了这样一种假设: 原始 Markov 链是在一种状态中准备的, 这种状态并不依赖于 $\mathcal{I}_y$ 上状态空间的部分; 如果这种情况不存在, 仍然可以进行类似的分析, 但是首项必须包括超过 $O(\varepsilon)$ 阶次的原始层, 在这种阶次上, $v(t)$ 能把 $\mathcal{I}_x \times \mathcal{I}_y$ 上的一个函数调整为仅是 $\mathcal{I}_x$ 上的一个函数.

9.7 练　　习

1. 当 Q_0, Q_1 由 (9.5.1) 和 (9.5.2) 给出时, 找出 Markov 链的不变测度的多尺度展开式, 其中生成子

$$Q = \frac{1}{\varepsilon} Q_0 + Q_1.$$

2. 找到 (9.5.3) 给出的 $\bar{Q}_1$ 的不变测度, 并在第 1 题答案的启示下解释你的发现.

3. 考虑随机微分方程 (6.5.1), 设 u 是由一个二态 Markov 链控制的, 其态 $\{-1, +1\}$. 写出 (z, u) 中合成 Markov 链的生成子, 其中在这个假设上, Markov 链的生成子有如下形式:

$$L(z) = \begin{pmatrix} -\theta(z) & \theta(z) \\ \phi(z) & -\phi(z) \end{pmatrix}.$$

4. 考虑与第 3 题相同的背景, 但是二态 Markov 链具有生成子 $\frac{1}{\varepsilon} L(z)$, 并且 $L(z)$ 是由第 3 题给出的. 使用平均化方法找出当 $\varepsilon \to 0$ 时, 关于 z 的平均随机微分方程, 其中 u 是可去掉的.

5. 设 u 是一个二态时间连续 Markov 链, 并且具有第 4 题中的生成子. 考虑常数微分方程

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \lambda(u) z, \quad t \in [0, \infty).$$

设 $\lambda(-1) < 0$ 和 $\lambda(+1) > 0$, 使用多尺度分析方法确定 z 的轨迹在什么条件下不增长.

6. 设 u 是具有生成子 Q 的有限状态空间上的一个 Markov 链, 其中 Q 源于

$$Q = \frac{1}{\varepsilon} Q_0 + Q_1.$$

设 Q_i 是当 $i=0,1$ 时 Markov 链的生成子, 并且 Q 有一个二维零空间

$$\mathcal{N}(Q_0)=\operatorname{span}\{\phi_0,\phi_1\}.$$

推导一个二态 Markov 链, 使得它近似于在这个零空间上的动力系统.

第 10 章 常微分方程和随机微分方程的平均化

10.1 引　　言

这里运用第 9 章讲到的 Markov 链的平均化原理, 将其应用到常微分和随机微分方程中. 通过扰动展开和变量消除, 得到向后方程近似解的统一格式.

10.2 节中给出了将要研究的方程; 10.3 节给出了平均化方程; 10.4 节中给出了平均化方程的推导; 10.5 节中研究了怎样处理确定性条件下的情形; 10.6 节对了两个例子进行了说明; 10.7 节把本章中的结论进行了推广, 并给出了一些参考评论.

10.2 完 全 方 程

若 $z=\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}$, 并且考虑如下情况:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}=f\left(x,y\right), \quad x\left(0\right)=x_0, & (10.2.1\mathrm{a}) \\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=\dfrac{1}{\varepsilon}g\left(x,y\right)+\dfrac{1}{\sqrt{\varepsilon}}\beta\left(x,y\right)\dfrac{\mathrm{d}V}{\mathrm{d}t}, \quad y\left(0\right)=y_0, & (10.2.1\mathrm{b}) \end{cases}$$

则 z 是方程 (6.1.1) 的解, 其中 $\varepsilon\ll 1$ 且 V 是一个标准布朗运动, $x\in\mathcal{X}$, $y\in\mathcal{Y}$, $z\in\mathcal{Z}$, 其符号如 4.1 节和 6.1 节所示.

在第 8 章中考虑了这样一个系统: 在这个系统中, 快速收敛于一个与 x 相关的固定点. 这就引出一种情形: 变量 y 是随着变量 x 变化的. 平均化将这种观点推广到另外一种情况: 当 x 是固定时, 变量 y 的动态性将会更复杂. 正如第 9 章中关于 Markov 链的处理方式, 可以通过一种适当的不变测度对快速变量 y 求平均. 现在为了使这种直接推测更准确, 定义生成子

$$\mathcal{L}_0=g\left(x,y\right)\cdot\nabla_y+\frac{1}{2}B\left(x,y\right):\nabla_y\nabla_y, \tag{10.2.2a}$$

$$\mathcal{L}_1=f\left(x,y\right)\cdot\nabla_x, \tag{10.2.2b}$$

其中 $B\left(x,y\right)=\beta\left(x,y\right)\beta\left(x,y\right)^{\mathrm{T}}$. 在本节中, 为了实现平均化过程, 作出一个遍历性假设. 最有效的方法就是假定对任意固定的 x, $\mathcal{L}_0$ 都有一个一维零空间, 由

$$\mathcal{L}_0 1\left(y\right)=0, \tag{10.2.3a}$$

$$\mathcal{L}_0^*\rho^\infty(y;x)=0 \tag{10.2.3b}$$

来表征, 其中 $1(y)$ 表示 y 里面的常数. 在 $\mathcal{Y}=\mathbb{T}^d$ 的情况下, 算子 $\mathcal{L}_0$ 和 $\mathcal{L}_0^*$ 都有周期边界条件. 在这种情况下, 如果 $B(x,y)$ 是严格正定的且在 $(x,y)\in\mathcal{X}\times\mathcal{Y}$ 上是一致的 (如定理 6.16 所示), 则关于算子 $\mathcal{L}_0$ 和 $\mathcal{L}_0^*$ 的零空间的假设被认为是成立的. 更一般的情况, 如当 $\mathcal{Y}=\mathbb{R}^d$ 或矩阵值函数 $B(x,y)$ 是退化的时, 相类似的严格证明是可能的, 但是在典型的情况下, 当采用带权的 L^p 空间时 (在这种空间中能体现无限远处不变密度的衰减特性), 这种泛函背景更为复杂, 见 18.4 节的备注.

10.3 简化方程

设快速过程 $y(t)$ 的生成子 $\mathcal{L}_0$, 对于任意 $x\in\mathcal{X}$ 满足 (10.3.1). 由

$$F(x)=\int_{\mathcal{Y}}f(x,y)\,\mu_x(\mathrm{d}y) \tag{10.3.1}$$

来定义向量场 F, 其中 $\mu_x(\mathrm{d}y)=\rho^\infty(y;x)\,\mathrm{d}y$.

结论 10.1 对于 $\varepsilon\ll 1$ 以及与 $O(1)$ 同阶的 t, X 近似于 (10.2.1) 的解 $x(t)$, 其中 X 是如下方程的解:

$$\frac{\mathrm{d}X}{\mathrm{d}t}=F(x),\quad X(0)=x_0. \tag{10.3.2}$$

注 10.2 若慢变量 x 是随机的且有如下形式:

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x,y)+\alpha(x,y)\frac{\mathrm{d}U}{\mathrm{d}t},\quad x(0)=x,$$

则类似的结论也成立, 其中 U 是独立于 V 的标准布朗运动. 在结论 10.1 成立的情况下, 平均化方程变为

$$\frac{\mathrm{d}X}{\mathrm{d}t}=F(X)+A(X)\frac{\mathrm{d}U}{\mathrm{d}t},\quad X(0)=x,$$

其中 $F(X)$ 是相同的且

$$A(X)A(X)^{\mathrm{T}}=\int_{\mathcal{Y}}\alpha(x,y)\alpha(x,y)^{\mathrm{T}}\mu_x(\mathrm{d}y).$$

见练习第 1 题. □

10.4 推　导

正如对 Markov 链, 采用向后 Kolmogorov 方程推导均匀化方程的. 设

$$v(x,y,t)=\mathbb{E}\left(\phi(x(t),y(t))\,|x(0)=x,y(0)=y\right).$$

随机偏微分方程 (10.2.1) 的向后方程 (6.3.4) 是

$$\frac{\partial v}{\partial t}=\frac{1}{\varepsilon}\mathcal{L}_0 v+\mathcal{L}_1 v, \tag{10.4.1}$$

其中 $\mathcal{L}_0, \mathcal{L}_1$ 是由 (10.2.2) 给出的, (6.3.4) 中的 z 是这里的 (x, y). $\mathcal{L}_0$ 是 y 上的一个微分算子, x 是其中的一个参量, 因此, 必须给它附加一个边界条件. 简单地假设给定了一个合适的边界条件, (10.2.3) 成立. 在这个例子中, 设 $\mathcal{Y}=\mathbb{T}^d$, 运用周期边界条件, 采用第 7 章的严密结论和快速过程的遍历性假设. 然而, 针对于其他泛函背景也是可能的, 接下来的关键是把 Fredholm 性质应用到由 $\mathcal{L}_0$ 定义的算子方程.

为了寻找在多尺度形式

$$v=v_0+\varepsilon v_1+O\left(\varepsilon^2\right)$$

下 (10.4.1) 的一个解, 可得

$$O\left(\frac{1}{\varepsilon}\right): \mathcal{L}_0 v_0=0; \tag{10.4.2a}$$

$$O(1): \mathcal{L}_0 v_1=-\mathcal{L}_1 v_0+\frac{\partial v_0}{\partial t}. \tag{10.4.2b}$$

(10.4.2a) 暗含了 v_0 属于 $\mathcal{L}_0$ 的零空间, 因此, 由 (10.2.3) 及遍历性可知其仅是 (x, t) 的函数. 固定 x, 则 (10.4.2b) 中的 Fredholm 性质被看成是 y 上的一个微分方程, 可得

$$-\mathcal{L}_1 v_0+\frac{\partial v_0}{\partial t} \perp \operatorname{Null}\left\{\mathcal{L}_0^*\right\}.$$

由 (10.2.3) 可知上式暗含了

$$\int_{\mathcal{Y}} \rho^{\infty}(y; x)\left(\frac{\partial v_0}{\partial t}(x, t)-f(x, y) \cdot \nabla_x v_0(x, t)\right) \mathrm{d} y=0.$$

由 ρ^{∞} 是一个概率密度可知 $\int_{\mathcal{Y}} \rho^{\infty}(y; x) \mathrm{d} y=1$. 因此,

$$\frac{\partial v_0}{\partial t}-\left(\int_{\mathcal{Y}} f(x, y) \mu_x(y) \mathrm{d} y\right) \cdot \nabla_x v_0(x, t)=0,$$

从而由 (10.3.1) 可得

$$\frac{\partial v_0}{\partial t}-F(x) \cdot \nabla_x v_0=0.$$

这是 (10.3.2) 的向后方程. 实际上, 定理 4.6 给出的特征线法表明已经得到了需要的结论.

10.5　确定性问题

本节提出了一个关于平均化方程的观点, 它是非常有用的, 主要有两个原因:

当 (10.2.1) 是确定性时, 也是适用的; 它形成了无论是确定性还是随机性背景下计算有效方程数值方法的基础. 切入点就是将 x 作为固定参数来分析 y 的快速动力系统的特性.

设 $\varphi_x^t(y)$ 是快速动力系统的一个解算子, 其中 x 是固定参数且 $\varepsilon = 1$. 更确切地说, 对于固定的 ξ 有

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_\xi^t(y) = g\left(\xi, \varphi_\xi^t(y)\right) + \beta\left(\xi, \varphi_\xi^t(y)\right)\frac{\mathrm{d}V}{\mathrm{d}t}, \quad \varphi_\xi^0(y) = y. \tag{10.5.1}$$

如第 8 章中 (10.2.1b) 的解 $y(t)$, 当 t 是 $o(1)$ (t 是常数的无穷小量), 并且它的近似值可以由 $y(t) \approx \varphi_{x(0)}^{t/\varepsilon}(y)$ 给出, 由于 t 是无穷小量, 从而使得 x 没有大的变化. 假定 (10.5.1) 是具有不变测度 μ_ξ 的遍历性, 在时间尺度上与 1 相比非常小, 而与 ε 相比非常大, 希望 $x(t)$ 近似固定, 并且由于 $y(t)$ 变化非常迅速, 因此, 也希望 $y(t)$ 在这个时间尺度上横跨自身 (依赖于 x) 的不变测度, 从而沿着 (10.5.1) 中的不变测度 ($\xi = x(t)$), 可以很自然地在 $x(t)$ 方程上将 $y(t)$ 平均化.

当 $\beta \equiv 0$ 时, $\varphi_\xi^t(y)$ 与 (8.2.2) 的解相符; 当 $\beta \neq 0$ 时, $\varphi_\xi^t(y)$ 依赖于布朗运动 $\{V(s)\}_{s\in[0,t]}$, 因此, 是一个随机过程. 这里设 $\varphi_\xi^t(y)$ 是遍历性的 (见 6.4 节), 而不是 (8.2.3) 中假设收敛于一个固定点. 这暗含了由

$$\mu_x(A) = \lim_{T\to\infty}\frac{1}{T}\int_0^T I_A\left(\varphi_x^t(y)\right)\mathrm{d}t, \quad A \subseteq \mathbb{T}^d \tag{10.5.2}$$

定义的测度对于任意 Borel 集 $A \subseteq \mathcal{Y}$ 的简单函数 I_A 都存在. (10.3.1) 中的平均化向量场 F 可以用这个测度来定义.

当研究随机微分方程 ($\beta \neq 0$) 时, 很自然地假设 $\mu_x(\cdot)$ 有一个与 Lebesgue 测度相关的密度, 使得 $\mu_x(\mathrm{d}y) = \rho^\infty(y;x)\,\mathrm{d}y$. 事实上, 对系数 $g(x,y)$ 和 $\beta(x,y)$ 给定适当的假设, 就可以证明这样一个密度的存在. 然而, 也可以通过一个例子说明在 Hamiltonian 原理中, 这个假设是不必要的. 注意到在第 8 章中, 测度 $\mu_x(\mathrm{d}y)$ 是一个 Dirac 质量, 它表征了不变流形 $\mu_x(\mathrm{d}y) = \delta(y - \eta(x))\,\mathrm{d}y$. 在这种情况下可得

$$F(x) = f(x, \eta(x)).$$

这是 (8.3.2) 中一个较为精确的向量场, 所以第 8 章中的简化方程是这里推导的一种特殊情形. 但是当 β 为非零时, 并设测度 μ 含有一个与 Lebesgue 测度相关的光滑密度 $\rho^\infty(y;x)$(这就是说, 假定 (10.2.3) 成立且 $\mu_x(\mathrm{d}y) = \rho^\infty(y;x)\,\mathrm{d}y$), 就可以推导出结论 10.1. 有一个平均化方程的表达式是很有用的, 因为它对确定性问题以及确定性或随机性问题中 F 的数值构造也是非常有效的. 下面通过时间平均表示遍历性平均来完成这样一个过程.

结论 10.3 通过时间平均

$$F(x)=\lim_{T\to\infty}\frac{1}{T}\int_0^T f(x,\varphi_x^s(y))\mathrm{d}s \tag{10.5.3}$$

表示 $F(x)$ 的另外一种表示形式, 运用 (10.5.2) 来评估 (10.3.1) 可以得到这个表示. 注意: 由遍历性所得到的平均并不依赖于 y.

10.6 应　用

考虑平均化原理的两个应用, 第一个是在随机微分方程的背景中, 第二个是在 Hamiltonian 常微分方程背景中.

10.6.1 不对称积随机微分方程

考虑方程

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=\left(1-y^2\right)x,\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=-\dfrac{\alpha}{\varepsilon}y+\sqrt{\dfrac{2\lambda}{\varepsilon}}\frac{\mathrm{d}V}{\mathrm{d}t},\end{cases}$$

其中 $\mathcal{X}=\mathcal{Y}=\mathbb{R}$. 寻找 x 是随时间增长还是否保持有界, 这是一个很有趣的问题. 当 $\varepsilon\to 0$ 时, 通过推导平均化方程可以了解这个问题. 从例 6.19 可知 y 是 OU 过程的一个时间尺度. 遍历性过程 y 的不变测度满足零期望值的 Gaussian 分布 $\mathcal{N}\left(0,\dfrac{\lambda}{\alpha}\right)$(见例 6.19). 注意: 这个测度并不依赖于 x, 因此, 仅有一个密度 $\rho^\infty(y)$. 在这里, 由

$$F(x)=\left(1-\int_{\mathbb{R}}\rho^\infty(y)\,y^2\mathrm{d}y\right)x$$

来定义平均化向量场 F, 其中 ρ^∞ 是与 $\mathcal{N}(0,\lambda/\alpha)$ 相关的密度, 从而

$$\int_{\mathbb{R}^d}\rho^\infty(y)\,y^2\mathrm{d}y=\frac{\lambda}{\alpha}$$

和

$$F(x)=\left(1-\frac{\lambda}{\alpha}\right)x.$$

因此, 平均化方程是

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\left(1-\frac{\lambda}{\alpha}\right)X.$$

从以上也可以推导出: 如果 $\lambda=\alpha$, 则 x 的轨迹趋于无限; 若 $\lambda>\alpha$, 其轨迹为收缩的; 如果 $\lambda=\alpha$, 则平均化向量场为零. 本节中需要重新给定时间尺度 $t\mapsto t/\varepsilon$, 以

得到问题

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon}\left(1-y^2\right)x, \\ \dfrac{\mathrm{d}x}{\mathrm{d}t} = -\dfrac{\alpha}{\varepsilon^2}y + \sqrt{\dfrac{2\alpha}{\varepsilon^2}}\dfrac{\mathrm{d}v}{\mathrm{d}t}. \end{cases}$$

在这个更长的时间尺度上, 将发生非平凡动力系统. 这种形式的随机微分方程是第 11 章的主题, 其具体的例子见 11.7 节.

10.6.2 Hamiltonian 原理①

在许多应用中, 将会遇到具有强势外力的 Hamiltonian 系统, 这些强势外力表现为一个固定子流形周围的快速小幅振荡. 通过快速振动变量的平均化来描述缓慢变化自由度的进化过程是一个有趣的思想. 现在给出这样的一个例子. 这个例子也是很有趣的, 它说明本章中的形式论可以被拓展用来研究很一般的无噪声微分方程, 也说明了处理极限措施 μ 保留一些初始条件 (系统总能量) 记忆的情况也是可能的.

考虑一个双粒子的 Hamiltonian 系统

$$H(x,p,y,v) = \frac{1}{2}\left(p^2+v^2\right) + \Phi(x) + \frac{\omega(x)}{2\varepsilon^2}y^2, \tag{10.6.1}$$

其中 (x,y) 是坐标系, (p,v) 是双粒子的共轭矩, $\Phi(x)$ 是一个非负势, 并假设对于所有的 x, $\omega(x)$ 满足 $\omega(x) \geqslant \varpi > 0$. 相应的运动方程为

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = p, \\ \dfrac{\mathrm{d}p}{\mathrm{d}t} = -\Phi'(x) - \dfrac{\omega'(x)}{2\varepsilon^2}y^2, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = v, \\ \dfrac{\mathrm{d}v}{\mathrm{d}t} = -\dfrac{\omega(x)}{\varepsilon^2}y. \end{cases}$$

设 E 表示 H 在时间 $t=0$ 时的 Hamiltonian 值,

$$E = H\left(x(0), p(0), y(0), v(0)\right),$$

E 是双粒子系统的总能量. 设 E 有界且独立于 ε, 由于 Hamiltonian 的 H 在时间上是守恒的, Φ 是非负的且 $\omega \geqslant \varpi$, 方程 (10.6.1) 暗含了

$$y^2 \leqslant \frac{2\varepsilon^2 E}{\varpi}.$$

① 这个实例与 R.Kupferman 协作而得到发展.

因此, 当 $\varepsilon \to 0$ 时, 其解就近似于 $y=0$ 的子流形. 然而, 要注意的是, y 出现在 x 方程里 y/ε 的组合以及对能量 H 的表达式中, 从而很自然地对变量 $\eta=\eta/\varepsilon$ 作了一个改变, 即方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}=p, \\ \dfrac{\mathrm{d}p}{\mathrm{d}t}=-\Phi'(x)-\dfrac{\omega'(x)}{2}\eta^2, \\ \dfrac{\mathrm{d}\eta}{\mathrm{d}t}=\dfrac{1}{\varepsilon}v, \\ \dfrac{\mathrm{d}v}{\mathrm{d}t}=-\dfrac{\omega(x)}{\varepsilon}\eta. \end{cases} \tag{10.6.2}$$

在这些变量中, 重新获得一个具有慢变量 $x\leftarrow(x,p)$, 快变量 $y\leftarrow(\eta,v)$ 和 (10.2.1) 形式的系统. 把上述方程写成二阶的形式为

$$\begin{cases} \dfrac{\mathrm{d}^2x}{\mathrm{d}t^2}+\Phi'(x)+\dfrac{1}{2}\omega'(x)\eta^2=0, \\ \dfrac{\mathrm{d}^2\eta}{\mathrm{d}t^2}+\dfrac{1}{\varepsilon^2}\omega(x)\eta=0. \end{cases}$$

此二阶形式的方程具有指导意义. 快速方程表示一个简谐振子, 这个简谐振子的频率 $\omega^{1/2}(x)$ 由变量 x 控制.

考虑快速动力系统, 其中 (x,p) 是固定的. 当 $\varepsilon=1$ 及 x 固定时, 这种快速动态的 Hamiltonian 为

$$H_{\text{fast}}=\frac{1}{2}v^2+\frac{\omega(x)}{2}\eta^2.$$

在给定的 (x,p) 中, 当 (x,p) 固定时, 快速系统的能量是守恒的, 并且可以由初始化系统的总能量减去与固定变量相关的能量来得到. 用

$$E_{\text{fast}}=E-\frac{1}{2}p^2-\Phi(x)$$

表示这个计算的结果. 对于固定的 x,p, η,v 中的动力性受限于能壳 $H_{\text{fast}}(v,\eta)=E_{\text{fast}}$. 用 $\mathcal{Y}(x,p)$ 表示这个能壳, 它是通过固定变量 (x,p) 来参数化的.

对谐振子的研究见例 4.17. 同理, 可推出采用 $\mathcal{Y}(x,p)$ 上遍历性测度 $\mu_{x,p}$ 而得到的快速振动的动能平均是

$$\int_{\mathcal{Y}(x,p)}\frac{\omega(x)}{2}\eta^2\mu_{x,p}(\mathrm{d}\eta,\mathrm{d}v)=\frac{1}{2}\left[E-\frac{1}{2}p^2-\Phi(x)\right],$$

从而

$$\int_{\mathcal{Y}(x,p)}\frac{1}{2}\eta^2\mu_{x,p}(\mathrm{d}\eta,\mathrm{d}v)=\frac{1}{2\omega(x)}\left[E-\frac{1}{2}p^2-\Phi(x)\right],$$

其中 (x,p) 被看成是固定参数, 并且总能量 E 是由整个系统的初值确定的. 平均化原理表明关于 p 的方程 (10.6.2) 中的快速变化量 η^2 能够被它的遍历性平均值近似, 当给定 $(X,P)\approx(x,p)$ 时, 就会产生一个封闭的方程系统, 它们是

$$\begin{cases} \dfrac{\mathrm{d}X}{\mathrm{d}t} = P, \\ \dfrac{\mathrm{d}P}{\mathrm{d}t} = -\Phi'(X) - \dfrac{\omega'(X)}{2\omega(X)}\left[E - \dfrac{1}{2}P^2 - \Phi(x)\right], \end{cases} \tag{10.6.3}$$

其中初值 E, $X(0) = X_0 = x(0)$ 以及 $P(0) = P_0 = p(0)$. 将在下面证明满足 (10.6.3) 的 (X,P) 保持如下绝热不变量:

$$J = \frac{1}{\omega^{1/2}(X)}\left[E - \frac{1}{2}P^2 - \Phi(X)\right],$$

从而 (10.6.3) 减少到 Hamiltonian 形式

$$\frac{\mathrm{d}X}{\mathrm{d}t} = P, \tag{10.6.4a}$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\Phi'(X) - J_0\left[\omega^{1/2}(X)\right]', \tag{10.6.4b}$$

其中 J_0 由下式给出:

$$J_0 = \frac{1}{\omega^{1/2}(X_0)}\left[E - \frac{1}{2}P_0^2 - \Phi(X_0)\right].$$

这意味着依赖于慢变量的刚性势的影响可以导出一个 Hamiltonian 结构, 可由一个有效势

$$\Phi_{\mathrm{eff}}(x) = \Phi(x) + J_0\omega^{1/2}(x)$$

代替势 $\Phi(x)$. 极限方程通过常数 J_0 包含了快速变量初始条件的记忆, 从而这种情况略不同于由结论 10.1 和结论 10.3 综合所覆盖的情况.

注意到可以通过 J 的定义以及 (10.6.3) 证明 J 在时间上实际是守恒的,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\omega^{1/2}(X)J\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(E - \frac{1}{2}P^2 - \Phi(X)\right) \\ &= -P\frac{\mathrm{d}P}{\mathrm{d}t} - \Phi'(X)\frac{\mathrm{d}X}{\mathrm{d}t} \\ &= \frac{P\omega'(X)}{2\omega(X)}\left(E - \frac{1}{2}P^2 - \Phi(X)\right) \\ &= \frac{P\omega'(X)}{2\omega^{1/2}(X)}J. \end{aligned}$$

但是, 由于 $\mathrm{d}X/\mathrm{d}t = P$, 于是可以得到另外的表达式

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\omega^{1/2}(X)J\right) &= \frac{1}{2}\frac{\omega'(X)}{\omega^{1/2}(X)}\frac{\mathrm{d}X}{\mathrm{d}t}J + \omega^{1/2}(X)\frac{\mathrm{d}J}{\mathrm{d}t}\\ &= \frac{P\omega'(X)}{2\omega^{1/2}(X)}J + \omega^{1/2}(X)\frac{\mathrm{d}J}{\mathrm{d}t}.\end{aligned}$$

由于 $\omega(X)$ 是严格正的, 使两个表达式相等可得

$$\frac{\mathrm{d}J}{\mathrm{d}t} = 0.$$

10.7 讨论和参考

平均化是基于某些快速过程的形式或遍历性的; 这个过程是确定的或者随机的并不重要. 然而, 一般来说, 建立随机问题的遍历性更简单些, 这就是普通发展受限于这种情况的原因. 应用于 (10.2.1) 的平均化方法是在启发性方式下来分析的, 见文献 [240], 是采用 Liouville 方程来构造平均极限的一个严格证明. 有时, 当快速过程的零空间是有限维而不是一维时, 得到无遍历性下的平均化结论也是可能的, 见文献 [246, 326].

常微分方程平均化方法的详细叙述以及大量例题可见文献 [281], 也可见文献 [13]. 俄语文献的英文综述可见文献 [281]. 关于慢流形主题的综述, 尤其是在 Hamiltonian 问题的背景下的综述可见文献 [193]. [321] 提供了一个对大量具有尺度分离问题下使用消元法的综述.

Anosov's 定理是常微分方程 (10.2.1)($\beta \equiv 0$) 背景下平均化原理的名称. 这个定理要求快速动力系统是具有遍历性的. 通常遍历性失效的原因是 "共振区"(在 $\mathcal{X}$ 上且快速动态不是遍历性的一个区域) 的存在. Arnld 和 Neistadt[193] 将 Anosov's 结论拓展到另一种情况. 在这种情况下, 当 $x \in \mathcal{X}$ 集合充分小时, 遍历性假设失效. 这些结论由 Kifer 进一步推广和拓展到随机框架中, Kifer 也研究了扩散以及有效解和精确解之间差异性的偏差特性, 见文献 [169~172], 也可见文献 [111, 第 7 章].

快速动力系统趋于固定点、周期解或者混沌解的情况, 可以通过对 Young 测度的说明, 用一个统一的方式来处理, 见文献 [29, 309]. Artstein 和合作者考虑到关于 (10.2.1) 形式的一类奇异扰动系统, 并注意快变量和慢变量的极限性质. 在所有的这些情况中, 可以看出 (x, y) 收敛于 (X, μ_X), 其中 X 是

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \int_{\mathbb{T}^d} f(X, y)\mu_X(\mathrm{d}y)$$

的解, μ_X 是 $\mathbb{T}^d$ 上的遍历性测度; 在 Young 测度意义上 y 收敛于 μ_X. (在固定一个

点的情况下, Young 测度是集中在一点上的 Dirac 质量.) 沿着这些线一个通用定理的证明见文献 [17].

在随机微分方程背景下, (10.2.1) 的一个有趣的推广形式为

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x,y)+\alpha(x,y)\frac{\mathrm{d}U}{\mathrm{d}t},\tag{10.7.1a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{\varepsilon}g(x,y)+\frac{1}{\sqrt{\varepsilon}}\beta(x,y)\frac{\mathrm{d}V}{\mathrm{d}t}.\tag{10.7.1b}$$

从而简化方程是一个随机微分方程, 而不是常微分方程 (见注 10.2), 这种情况是第 11 章中考虑到的背景的子情况, 这可以通过令 $f_0=0$, 设 $f_1=f$, 并识别这里的 ε 和 ε^2 得到.

在应用章节中研究了双尺度 Hamiltonian 系统的平均化原理. 具有两个时间尺度的 Hamiltonian 问题 (由 Rubin 和 Ungar 发起) 的系统化研究见文献 [277]. 更多基于规范性理论最新的 Neistadt 观点, 已经被应用于这类问题[32]; 这种方法功能强大, 可以导出初始变量和极限变量之间的指数型误差估计. 处理这个问题的另一个不同方法, 就是运用时间的均匀化技巧[43], 即论文 [44]. 10.6.2 小节中介绍的例子就是源于那篇论文. 这里给出的启发式推导, 运用时间均匀化方法的严密证明见 [44], 并且它也可以被推广到更高的维数. 随着 m 的增加, 限制了平均化方法在这类双尺度 Hamiltonian 系统中的应用, 共振变得越来越重要 (源于文献 [306]).

第 11 章将对多尺度常微分方程和随机微分方程的数值工作进行阐述.

10.8 练　　习

1. 假设 U 和 V 是相互独立的标准布朗运动, 从随机微分方程 (10.7.1) 中推导出平均化方程 (见注 10.2).

2. 设 $\Phi:\mathcal{X}\times\mathcal{Y}:\mathbb{R}^+$, 并考虑方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=-\nabla_x\Phi(x,y)+\sqrt{2\sigma}\frac{\mathrm{d}U}{\mathrm{d}t},$$
$$\frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{1}{\varepsilon}\nabla_y\Phi(x,y)+\sqrt{\frac{2\sigma}{\varepsilon}}\frac{\mathrm{d}V}{\mathrm{d}t},$$

其中 U 和 V 是适当维数的标准布朗运动. 在一个 Fredholm 性质假设下, 清晰的陈述并说明 X 的平均化方程具有如下形式:

$$\frac{\mathrm{d}X}{\mathrm{d}t}=-\nabla\Psi(X)+\sqrt{2\sigma}\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中 Ψ 是由

$$\exp\left(-\frac{1}{\sigma}\Psi(x)\right)=\int_{\mathcal{Y}}\exp\left(-\frac{1}{\sigma}\Phi(x,y)\right)\mathrm{d}y$$

给出的, W 是适当维数的布朗运动. (事实上, 强收敛方法, 如 17 章中重点提出的这些, 可以用来说明当 $W=U$ 时, $X\approx x$.)

3. 设 $\varPhi$ 如第 2 题中给出的, 将如下二阶系统写成耦合一阶随机微分方程:

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+\frac{\mathrm{d}x}{\mathrm{d}t}=-\nabla_x\varPhi(x,y)+\sqrt{2\sigma}\frac{\mathrm{d}U}{\mathrm{d}t},$$

$$\varepsilon\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+\frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{1}{\varepsilon}\nabla_y\varPhi(x,y)+\sqrt{\frac{2\sigma}{\varepsilon}}\frac{\mathrm{d}V}{\mathrm{d}t}.$$

找出快速过程 y 的稳态分布, 并在第 2 题结论的指导下找出 X 的平均化方程.

4. 运用式 (10.5.3), 从 10.6.1 小节的例子中推导出平均化方程.

5. 设 u 是一个时间连续的 Markov 链, 具有生成子

$$L=\begin{pmatrix}-a & a\\ b & -b\end{pmatrix}$$

不失共性地标记状态空间 $\mathcal{I}=\{-1,+1\}$. 由 $\omega(\pm1)=\omega^{\pm}$ 和 $m(\pm1)=m^{\pm}$ 定义两个函数 $\omega:\mathcal{I}\to(0,\infty)$ 和 $m:\mathcal{I}\to(-\infty,\infty)$, 接下来考虑随机微分方程, 其系数取决于 u, 并且由

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x,y)+\sqrt{2\sigma}\frac{\mathrm{d}U}{\mathrm{d}t},$$

$$\frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{1}{\varepsilon}\omega(u)(y-m(u))+\sqrt{\frac{2\sigma}{\varepsilon}}\frac{\mathrm{d}V}{\mathrm{d}t}$$

给出, 其中 U 和 V 是适当维数的标准布朗运动. 写出过程 (x,y,u) 的生成子, 并运用多尺度分析推导出平均化耦合 Markov 链的形式为

$$\frac{\mathrm{d}X}{\mathrm{d}t}=F(X,u)+\sqrt{2\sigma}\frac{\mathrm{d}W}{\mathrm{d}t}$$

的随机微分方程, 其中 W 是与 U 具有相同维数的标准布朗运动.

6. 将第 5 题扩展到这样一种情况, 即 Markov 链的转移率是由 a 和 b 确定的, 取决于 x 和 y.

7. 运用时间平均化, 找出注 10.2 中有效系数矩阵 $A(x)$ 的一个替代.

第11章　常微分方程、随机微分方程的均匀化

11.1　引　　言

本章将继续研究具有两个广泛分离特征时间尺度的随机微分方程系统. 类似于前面章节, 所不同的是, 在本章中要导出一个在更长的扩散时间尺度上描述动力系统的有效方程. 例如, 由于问题的对称性, 当定义在方程 (10.3.1) 中的有效漂移项 $F(x)$ 消失时, 这就是一个有趣的时间尺度. 在中心条件下, 能够捕捉到这种消失的有效漂移, 见方程 (11.2.5). 与前面章节相比, 在扩散时间尺度下, 即使当噪声没有直接作用于慢变量上, 即在方程 (11.2.1) 中 $\alpha(x,y)\equiv 0$ 时, 有效方程也是随机的.

在 11.2 节中给出了本章分析的随机微分方程; 11.3 节包含了 11.4 节中推导出的简化方程; 11.5 节描述了简化方程的各种性质, 在推导时假设将要消失的快速过程是随机的; 11.6 节中介绍了如何处理确定性例子; 11.7 节中介绍了本章理论的各种应用: 11.7.1 小节和 11.7.2 小节分别给出了快速过程为 Ornstein-Uhlenbeck 过程和混沌决定性过程的例子, 推导 Stratonovich 随机积分, 并将其作为白噪声近似值的极限见 11.7.3 小节, 11.7.4 小节中研究了 Stokes 定律, 11.7.5 小节中推导了统计力学中的 Green-Kubo 公式, 在 11.7.6 小节中可知 Itô 方程和 Stratonovich 积分都不可以解释极限方程中的随机积分, 11.7.7 小节中研究了 Lévy 面积修正; 对本章结论的延伸及其参考文献见 11.8 节.

11.2　完 全 方 程

考虑如下随机微分方程:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}=\dfrac{1}{\varepsilon}f_0(x,y)+f_1(x,y)+\alpha(x,y)\dfrac{\mathrm{d}U}{\mathrm{d}t}, & x(0)=x_0, \qquad (11.2.1\mathrm{a})\\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t}=\dfrac{1}{\varepsilon^2}g(x,y)+\dfrac{1}{\varepsilon}\beta(x,y)\dfrac{\mathrm{d}V}{\mathrm{d}t}, & y(0)=y_0, \qquad (11.2.1\mathrm{b})\end{cases}$$

其中 U 和 V 为独立的标准 Brownian 运动. x 和 y 方程中都包含了快速动力系统, 但是 y 中动力系统的数量级要高于 x 中动力系统的数量级, 正如在 4.1 节及 6.1 节中讨论的一样, $x\in\mathcal{X}, y\in\mathcal{Y}$, 并且 $\mathcal{X}\oplus\mathcal{Y}=\mathcal{Z}$.

对于方程 (11.2.1), 向后 Kolmogorov 方程 (6.3.4) 如下①, 其中 $\phi = \phi(x)$:

$$\frac{\partial v}{\partial t} = \frac{1}{\varepsilon^2}\mathcal{L}_0 v + \frac{1}{\varepsilon}\mathcal{L}_1 v + \mathcal{L}_2 v, \quad (x,y,t) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^+, \tag{11.2.2a}$$

$$v = \phi(x), \quad (x,y,t) \in \mathcal{X} \times \mathcal{Y} \times \{0\}, \tag{11.2.2b}$$

其中

$$\mathcal{L}_0 = g \cdot \nabla_y + \frac{1}{2}B : \nabla_y \nabla_y, \tag{11.2.3a}$$

$$\mathcal{L}_1 = f_0 \cdot \nabla_x, \tag{11.2.3b}$$

$$\mathcal{L}_2 = f_1 \cdot \nabla_x + \frac{1}{2}A : \nabla_x \nabla_x, \tag{11.2.3c}$$

$$A(x,y) := \alpha(x,y)\alpha(x,y)^{\mathrm{T}},$$

$$B(x,y) := \beta(x,y)\beta(x,y)^{\mathrm{T}}.$$

在这个 Kolmogorov 方程中, 用多尺度方法消去 y, 从而确保得到一个仅含 x 的动力系统简化方程.

在 y 中, 将生成子 $\mathcal{L}_0$ 视为一个微分算子, $\mathcal{L}_0$ 中的 x 是一个参数, 对于生成子 $\mathcal{L}_0$ 来说, 通常用于消除变量的遍历性假设表明 $\mathcal{L}_0$ 具有如下特征的一维零空间:

$$\mathcal{L}_0 1(y) = 0, \tag{11.2.4a}$$

$$\mathcal{L}_0^* \rho^\infty(y;x) = 0, \tag{11.2.4b}$$

其中 $1(y)$ 定义为在 y 上为常数, 并且 $\rho^\infty(y;x)$ 是一个遍历性测度 $\mu_x(\mathrm{d}y) = \rho^\infty(y;x)\mathrm{d}y$ 的密度, 还假定在此测度下, $f_0(x,y)$ 的平均值为零, 使得**中心条件**

$$\int_{\mathcal{Y}} f_0(x,y)\mu_x(\mathrm{d}y) = 0, \quad \forall x \in \mathcal{X} \tag{11.2.5}$$

成立, 那么当极限 $\varepsilon \to 0$ 时, 由于 x 方程中包含 f_0 项, 所以将会对 x 的近似方程引起一个 $O(1)$ 的有效漂移和噪声.

就像前面章节中描述的那样, 当 $\mathcal{Y} = \mathbb{T}^d$ 时, 生成子 $\mathcal{L}_0$ 和 $\mathcal{L}_0^*$ 具有周期边界条件. 假定当 $(x,y) \in \mathcal{X} \times \mathbb{T}^d$ 时, $B(x,y)$ 严格一致正定, 则定理 6.16 证明 $\mathcal{L}_0^*$ 的零空间是一维的. 通常, 当 $\mathcal{Y} = \mathbb{R}^d$ 或者 $B(x,y)$ 是退化时, 也会有类似严格的证明, 但是函数设置可能会更加复杂, 在加权空间 L^p 上的情形就比较典型, 该空间的特点是其不变密度在无限远处衰减.

①为了简化起见, 使向后 Kolmogorov 方程的初始条件独立于 y, 这是没有必要的, 见 11.8 节的讨论.

当 $\mathcal{Y}=\mathbb{T}^d$ 且 $B(x,y)$ 是严格正定的, 定理 7.9 同样适用, 并且对于有如下形式的 Poission 方程有一个可解性理论:

$$-\mathcal{L}_0\phi=h\ . \tag{11.2.6}$$

特别地, 方程有解的充要条件为对于快速过程 $\mu_x(\mathrm{d}y)$ 的不变测度, 方程 (11.2.6) 的右边是处于中心位置,

$$\int_{\mathbb{T}^d} h(x,y)\mu_x(\mathrm{d}y)=0,\quad \forall x\in\mathcal{X}. \tag{11.2.7}$$

当满足 (11.2.7) 时, (11.2.6) 的解在 $\mathcal{L}_0$ 的零空间上是唯一常数, 令

$$\int_{\mathbb{T}^d}\phi\,(x,y)\,\mu_x\,(\mathrm{d}y)=0,\quad \forall x\in\mathcal{X}.$$

可以确定这个常数. 通常, 当 $\mathcal{Y}=\mathbb{R}^d$ 或 $B(x,y)$ 衰减时, Poission 方程 (11.2.6) 解的存在唯一性问题就变得更加复杂. 然而, 当函数空间使得衰减性在无限空间上增强时, 可以得到类似的结论, 见注 11.9 以及 11.8 节中所列的参考文献.

11.3 简化方程

假定算子 $\mathcal{L}_0$ 满足 Fredholm 性质 (定理 2.42), 并且具有特征为 (11.2.4) 的一维零空间. **单胞问题**①定义如下:

$$-\mathcal{L}_0\Phi\,(x,y)=f_0\,(x,y)\,,\quad \int_{\mathcal{Y}}\Phi\,(x,y)\,\rho^{\infty}\,(y;x)\,\mathrm{d}y=0. \tag{11.3.1}$$

可将其视为在 y 上并含有参数 x 的偏微分方程. 由于 f_0 满足 (11.2.5), 则通过 Fredholm 性质得 (11.3.1) 有唯一解. 向量场 F 定义如下:

$$\begin{aligned}F(x)&=\int_{\mathcal{Y}}\left(f_1(x,y)+(\nabla_x\Phi(x,y))f_0(x,y)\right)\rho^{\infty}(y;x)\mathrm{d}y\\&=F_1(x)+F_0(x),\end{aligned} \tag{11.3.2}$$

并定义扩散矩阵 $A(x)$ 为

$$A(x)A(x)^{\mathrm{T}}=A_1(x)+\frac{1}{2}(A_0(x)+A_0(x)^{\mathrm{T}}), \tag{11.3.3}$$

其中

$$A_0(x):=2\int_{\mathcal{Y}}f_0(x,y)\otimes\Phi(x,y)\rho^{\infty}(y;x)\mathrm{d}y, \tag{11.3.4}$$

① 这里的单胞指的是周期单胞, 当 $\mathcal{Y}=\mathbb{T}^d$ 时, 其为快变量设置了尺度. 这个术语取自于偏微分方程的周期均匀化.

$$A_1(x) := \int_{\mathcal{Y}} A(x,y)\,\rho^\infty(y;x)\mathrm{d}y. \tag{11.3.5}$$

为了确保定义 $A(x)$, 有必要证明 $A_1(x)$ 的和以及 $A_0(x)$ 的对称部分是半正定的, 见 11.5 节.

结论 11.1 因为 $\varepsilon \ll 1$ 且 t 为 $O(1)$, 所以过程 $x(t)$ 可由过程 $X(t)$ 逼近, 其中 $x(t)$ 为 (11.2.1) 的解, $X(t)$ 为下面方程的解:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) + A(X)\frac{\mathrm{d}W}{\mathrm{d}t}, \quad X(0) = x_0. \tag{11.3.6}$$

注 11.2 注意到由 AA^{T} 的信息不足以唯一确定 A, 因此, 方程 (11.3.3) 也不能唯一确定极限随机微分方程 (11.3.6). 这是因为可能存在很多具有同样生成子的随机微分方程, 进而表明只有当概率测度为弱收敛时, 由 (11.3.6) 的解得到 (11.2.1) 的解的近似才是正确的, 见第 18 章. □

11.4 推 导

找一个 (11.2.2) 解的多尺度展开式, 形式如下:

$$v = v_0 + \varepsilon v_1 + \varepsilon^2 v_2 + \cdots, \tag{11.4.1}$$

其中 $v_j = v_j(x,y,t)$. 将展开式 (11.4.1) 代入 (11.2.2), 比较两边 ε 的幂得到如下方程:

$$O(1/\varepsilon^2): -\mathcal{L}_0 v_0 = 0; \tag{11.4.2a}$$

$$O(1/\varepsilon): -\mathcal{L}_0 v_1 = \mathcal{L}_1 v_0; \tag{11.4.2b}$$

$$O(1): -\mathcal{L}_0 v_2 = -\frac{\partial v_0}{\partial t} + \mathcal{L}_1 v_1 + \mathcal{L}_2 v_0. \tag{11.4.2c}$$

(11.2.4) 中的式 (11.4.2a) 暗含展开式的第一项独立于 y, $v_0 = v_0(x,t)$. 现在用方程 (11.4.2b) 进行推导. 假设 (11.2.5) 成立, 因为相对于 $\varphi_x^t(\cdot)$ 的不变测度, $f_0(x,y)$ 是中心的, 则满足可解性条件, 并且由 (11.2.3b) 得

$$\mathcal{L}_1 v_0 = f_0(x,y)\cdot\nabla_x v_0(x,t).$$

方程 (11.4.2b) 变为

$$-\mathcal{L}_0 v_1 = f_0(x,y)\cdot\nabla_x v_0(x,t). \tag{11.4.3}$$

由于 $\mathcal{L}_0$ 在 y 上仅是一个微分算子, x 是其中的一个参数, 则 (11.4.3) 的通解形式如下:

$$v_1(x,y,t) = \Phi(x,y)\cdot\nabla_x v_0(x,t) + \Phi_1(x,t). \tag{11.4.4}$$

函数 Φ_1 在后面推导时没有起到作用, 故设为零, 则用解 v_1 表示作用于 v_0 上的一个线性算子. 因为目的是要找到一个 v_0 的闭合方程, v_1 的这种形式很好地表示了方程的解. 在 (11.4.3) 中代入 v_1 后表明 Φ 是单胞问题 (11.3.1) 的解. 条件 (11.2.5) 确保单胞问题有解, 并且规范化条件保证了解是唯一的. 现在讨论 (11.4.2c), 发现其右边有如下形式:

$$-\left(\frac{\partial v_0}{\partial t} - \mathcal{L}_2 v_0 - \mathcal{L}_1(\Phi \cdot \nabla_x v_0)\right).$$

因此, 对每个固定的 x, (11.4.2c) 的可解性要求

$$\begin{aligned}\frac{\partial v_0}{\partial t} &= \int_{\mathcal{Y}} \rho^{\infty}(y;x)\mathcal{L}_2 v_0(x,t)\mathrm{d}y + \int_{\mathcal{Y}} \rho^{\infty}(y;x)\mathcal{L}_1(\Phi(x,y)\cdot\nabla_x v_0(x,t))\mathrm{d}y,\\ &= I_1 + I_2.\end{aligned} \tag{11.4.5}$$

分别考虑右边的两项. 第一项为

$$\begin{aligned}I_1 &= \int_{\mathcal{Y}} \rho^{\infty}(y;x)\left(f_1(x,y)\cdot\nabla_x + \frac{1}{2}A(x,y):\nabla_x\nabla_x\right)v_0(x,t)\mathrm{d}y\\ &= F_1(x)\cdot\nabla_x v_0(x,t) + \frac{1}{2}A_1(x):\nabla_x\nabla_x v_0(x,t).\end{aligned}$$

现在考虑第二项 I_2. 注意到

$$\mathcal{L}_1 = (\Phi\cdot\nabla_x v_0) = f_0\otimes\Phi : \nabla_x\nabla_x v_0 + (\nabla_x\Phi f_0)\cdot\nabla_x v_0,$$

因此, $I_2 = I_3 + I_4$, 其中

$$I_3 = \int_{\mathcal{Y}} \rho^{\infty}(y;x)(\nabla_x\Phi(x,y)f_0(x,y))\cdot\nabla_x v_0(x,t)\mathrm{d}y$$

且

$$I_4 = \int_{\mathcal{Y}} \rho^{\infty}(y;x)(f_0(x,y)\otimes\Phi(x,y):\nabla_x\nabla_x v_0(x,t))\mathrm{d}y,$$

则

$$I_2 = F_0(x)\cdot\nabla_x v_0(x,t) + \frac{1}{2}A_0(x):\nabla_x\nabla_x v_0(x,t).$$

合并 (11.4.5) 右边的简化形式, 由 (2.2.2) 可知, 因为只要求对 A_0 的对称部分与其他对称矩阵计算 Frobenius 内积, 于是有下面的表达式:

$$\frac{\partial v_0}{\partial t} = F(x)\cdot\nabla_x v_0 + \frac{1}{2}A(x)A(x)^{\mathrm{T}} : \nabla_x\nabla_x v_0.$$

这是一个与 (11.3.6) 给出的可约化动力系统相对应的向后方程.

11.5　简化方程的性质

如果由 (11.3.3), (11.3.5) 给出的 $A(x)A(x)^{\mathrm{T}}$ 是非负定的, 则就很好地定义了有效随机微分方程 (11.3.6), 现在给出证明.

定理 11.3　考虑如下情形: $\mathcal{Y}=\mathbb{T}^d$, 并且 $\mathcal{L}_0$ 具有周期边界条件, 则

$$\langle\xi, A_1(x)\xi+A_0(x)\xi\rangle \geqslant 0, \quad \forall x\in\mathcal{X}, \xi\in\mathbb{R}^l.$$

由于 $A(x)A(x)^{\mathrm{T}}$ 是非负定的, 从而由 (11.3.3) 就很好地定义了实值矩阵函数 $A(x)$.

证明　设 $\phi(x,y)=\xi\cdot\Phi(x,y)$, 则 ϕ 为下列方程的解:

$$-\mathcal{L}_0\phi=\xi\cdot f_0.$$

由定理 6.12 有

$$\begin{aligned}\langle\xi, A_1(x)\xi+A_0(x)\xi\rangle &= \int_{\mathcal{Y}}\left(|\alpha(x,y)^{\mathrm{T}}\xi|^2-2(\mathcal{L}_0\phi(x,y))\phi(x,y)\right)\rho^\infty(y;x)\mathrm{d}y\\ &= \int_{\mathcal{Y}}\left(|\alpha(x,y)^{\mathrm{T}}\xi|^2+|\beta(x,y)^{\mathrm{T}}\nabla_y\phi(x,y)|^2\right)\rho^\infty(y;x)\mathrm{d}y\\ &\geqslant 0,\end{aligned}$$

则

$$\begin{aligned}\langle\xi, AA^{\mathrm{T}}\xi\rangle &= \langle\xi, A_1\xi\rangle+\frac{1}{2}\langle\xi,(A_0+A_0^{\mathrm{T}})\xi\rangle\\ &= \langle\xi,(A_1+A_0)\xi\rangle\geqslant 0.\end{aligned}$$ □

下面有两个重要的注解.

注 11.4　和前面章节证明的方法类似, 用 (6.3.11) 代替 Dirichlet 形式, 表明

$$\frac{1}{2}(A_0(x)+A_0(x)^{\mathrm{T}})=\int_{\mathcal{Y}}(\nabla_y\Phi(x,y)\beta(x,y)\otimes\nabla_y\Phi(x,y)\beta(x,y))\rho^\infty(y;x)\mathrm{d}y. \quad \square \tag{11.5.1}$$

注 11.5　由注 6.13 可知, 除 $\mathcal{Y}=\mathbb{T}^d$ 外, 前面定理适用于其他情形.　□

11.6　确定性问题

正如前面章节所述, 为了数值目的以及确定性问题, 将有效方程用时间平均来表示是非常有用的. 为此, 下面给出 $A_0(x)$ 和 $F_0(x)$ 的第二种表示形式. 设 $\varphi_\xi^t(y)$ 为方程 (10.5.1) 的解, 根据初值和能够产生 Brownian 运动的标准独立 Wiener 测

度, 用 $\mu_x(\cdot)$ 形成乘积测度, 用 $\mathbb{E}^{\mu_x}$ 表示. 用这些记号进行时间积分来表示单胞问题的解, 从而得到下面的公式, 本节末给出其推导过程.

结论 11.6 下面在时间以及 $\mathbb{E}^{\mu_x}$ 上积分, 可以找到向量场 $F_0(x)$ 的替换形式以及扩散矩阵 $A_0(x)$,

$$A_0(x) = 2\int_0^\infty \mathbb{E}^{\mu_x}(f_0(x,y)\otimes f_0(x,\varphi_x^t(y)))\,\mathrm{d}t, \tag{11.6.1}$$

并且如果生成子 $\mathcal{L}_0$ 独立于 x, 则

$$F_0(x) = \int_0^\infty \mathbb{E}^{\mu_x}(\nabla_x f_0(x,\varphi_x^t(y))f_0(x,y))\,\mathrm{d}t. \tag{11.6.2}$$

由遍历性知所有这些表达式对任意 y 均成立.

本结论中, 在 t 上积分使得没有用单胞问题 Φ 的显式解就能表示有效方程, 而且为了更好地定义, 要求校正值衰减得足够快.

在前面的结论中, 用 s 上的时间平均来代替期望可以找到另一对 $F(x)$ 和 $A(x)A(x)^{\mathrm{T}}$ 的替换形式, 那时 A_0 和 F_0 的表达式中含有两个时间积分：在 s 上的积分是一个遍历性平均, 其是对于路径空间上不变测度平均的替换; 在时间 t 上的积分没有用单胞问题 Φ 的显示解去表示有效方程, 并且为了更好地定义, 要求校正值衰减得足够快. 事实上, 单胞问题 (11.3.1) 的适定性暗含了校正值的衰减特性.

结论 11.7 通过在时间上的积分, 可以找到向量场 F 的替换形式以及扩散矩阵 A,

$$F_1(x) = \lim_{T\to\infty}\frac{1}{T}\int_0^T f_1(x,\varphi_x^s(y))\mathrm{d}s,$$

$$A_1(x) = \lim_{T\to\infty}\frac{1}{T}\int_0^T A(x,\varphi_x^s(y))\mathrm{d}s$$

且

$$A_0(x) = 2\int_0^\infty\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T f_0(x,\varphi_x^s(y))\otimes f_0(x,\varphi_x^{t+s}(y))\mathrm{d}s\right)\mathrm{d}t, \tag{11.6.3}$$

其中 $\varphi_x^t(y)$ 是 (10.5.1) 的解, 而且如果生成子 $\mathcal{L}_0$ 独立于 x, 则

$$F_0(x) = \int_0^\infty\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T \nabla_x f_0(x,\varphi_x^{t+s}(y))f_0(x,\varphi_x^s(y))\mathrm{d}s\right)\mathrm{d}t.$$

由遍历性知所有这些表达式对任意 y 均成立.

结论 11.6 和结论 11.7 中推导了 $A_0(x)$ 和 $F_0(x)$ 的替换形式, 下面的结论对上述推导过程会很有用. 这些结论使用遍历性作为时间平均来表示单胞问题以及相关 Poisson 方程的解.

结论 11.8　设 $\mathcal{L}$ 为 $\mathcal{Y}$ 上遍历性 Markov 过程 $y(t)$ 的生成子, 其满足下面的随机微分方程:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = g(y) + \beta(y)\frac{\mathrm{d}V}{\mathrm{d}t}, \quad y(t) = y, \tag{11.6.4}$$

并且设 $\mu(\mathrm{d}y)$ 为一致不变测度. 假设 h 相对于 μ 是中心的, 即

$$\int_{\mathcal{Y}} h(y)\mu(\mathrm{d}y) = 0,$$

则 Poisson 方程

$$-\mathcal{L}f = h, \quad \int_{\mathcal{Y}} f(y)\mu(\mathrm{d}y) = 0$$

的解 $f(y)$ 的表达式为

$$f(y) = \int_0^\infty (\mathrm{e}^{\mathcal{L}t}h)(y)\mathrm{d}t. \tag{11.6.5}$$

证明　对 $f(y(t))$ 使用 Itô 公式得

$$\begin{aligned}
f(y(t)) - f(y) &= \int_0^t \mathcal{L}f(y(s))\,\mathrm{d}s + \int_0^t \langle \nabla_y f(y(s)), \beta(y(s))\mathrm{d}W(s)\rangle \\
&= \int_0^t -h(y(s))\,\mathrm{d}s + \int_0^t \langle \nabla_y f(y(s)), \beta(y(s))\mathrm{d}W(s)\rangle.
\end{aligned}$$

关于 Wiener 测度求期望并使用随机积分的鞅性质, 以及由 $\mathbb{E}h(y(s)|y(0)=y)$ 为向后 Kolmogorov 方程的解得

$$f(y) = \mathbb{E}\,f(y(t)) + \int_0^t (\mathrm{e}^{\mathcal{L}s}h)(y)\mathrm{d}s.$$

取极限 $t \to \infty$, 使用过程 $y(t)$ 的遍历性, 并由相对于密度为 $\rho^\infty(y;x)$ 的不变测度 $f(y)$ 是中心的性质可推得

$$\begin{aligned}
f(y) &= \lim_{t\to\infty} \mathbb{E}\,f(y(t)) + \int_0^\infty (\mathrm{e}^{\mathcal{L}t}h)(y)\mathrm{d}t \\
&= \int_{\mathcal{Y}} f(y)\mu(\mathrm{d}y) + \int_0^\infty (\mathrm{e}^{\mathcal{L}t}h)(y)\mathrm{d}t \\
&= \int_0^\infty (\mathrm{e}^{\mathcal{L}t}h)(y)\mathrm{d}t.
\end{aligned}$$

证毕.　□

注 11.9　当将前面的结论应用于关于 μ 具有中心性质的函数时, 至少在形式上可以记为

$$\mathcal{L}^{-1} = -\int_0^\infty \mathrm{e}^{\mathcal{L}t}\mathrm{d}t,$$

而且对于 (11.6.4) 中参数依赖于 x 的情形, 该结论同样是成立的. □

本节通过时间积分推导出 $A(x)$ 和 $F(x)$ 的替换形式, 该替换形式在结论 11.7 和结论 11.6 中给出. 应用时间平均等于关于 ρ^∞ 平均, 由遍历性可得结论 11.7 中 $F_1(x)$ 和 $A_1(x)$ 的表达式. 由结论 11.8, 可将单胞问题的解记为

$$\Phi(x,y) = \int_0^\infty (\mathrm{e}^{\mathcal{L}_0 t} f_0)(x,y)\mathrm{d}t = \int_0^\infty \mathbb{E}\, f_0(x, \varphi_x^t(y))\mathrm{d}t, \tag{11.6.6}$$

其中 $\mathbb{E}$ 定义为关于 Wiener 测度的期望, 则

$$F_0(x) = \int_{\mathcal{Y}} \rho^\infty(y;x)\nabla_x \Phi(x,y) f_0(x,y)\mathrm{d}y.$$

当 $\mathcal{L}_0$ 独立于 x 时, 使得 $\varphi_x^t(\cdot) = \varphi^t(\cdot)$ 也独立于 x, 就像 $\mu_x = \mu$ 和 $\rho^\infty(\cdot\,;x) = \rho^\infty(\cdot)$ 那样, 由 (11.6.6) 得

$$F_0(x) = \int_{\mathcal{Y}} \rho^\infty(y;x) \int_0^\infty \mathbb{E}\nabla_x f_0(x, \varphi^t(y)) f_0(x,y)\,\mathrm{d}t\mathrm{d}y,$$

其中 $\mathbb{E}$ 定义为关于 Wiener 测度的期望. 回想令 y 分布于其不变测度上得到乘积测度 $\mathbb{E}^{\mu_x}$, 以及产生 $\varphi^t(y)$ 方程的 Brownian 运动. 改变积分的阶数得

$$F_0(x) = \int_0^\infty \mathbb{E}^{\mu_x}(\nabla_x f_0(x, \varphi_x^t(y)) f_0(x,y))\,\mathrm{d}t. \tag{11.6.7}$$

这就是结论 11.6 要求得到的表达式. 用时间平均代替 $\mathbb{E}^{\mu_x}$ 上的平均, 对于所有的 y 有

$$F_0(x) = \int_0^\infty \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \nabla_x f_0(x, \varphi_x^{t+s}(y)) f_0(x, \varphi_x^s(y))\mathrm{d}s \right) \mathrm{d}t,$$

从而得到结论 11.7 中想要的公式.

类似于得到 (11.6.7) 的计算可以给出在结论 11.6 中 $A_0(x)$ 的公式 (11.6.1). 用时间平均代替 $\mathbb{E}^{\mu_x}$ 上的平均可得结论 11.7 中 $A_0(x)$ 的公式.

11.7 应　用

本节给出大量的例子来说明上述思想的广泛应用.

11.7.1 快速 Ornstein-Uhlenbeck 噪声

考虑方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{\varepsilon}(1-y^2)x, \tag{11.7.1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{\alpha}{\varepsilon^2}y + \sqrt{\frac{2\alpha}{\varepsilon^2}}\frac{\mathrm{d}V}{\mathrm{d}t}, \tag{11.7.2}$$

其中 $V(t)$ 是一维标准 Brownian 运动,

$$f_0(x,y) = (1-y^2)x, \quad f_1(x,y) = 0.$$

回想由例 6.19 可得 y 的方程是 OU 过程的一个时间尺度变换, 其中 $\lambda = \alpha$, 而且这些方程可由 10.6 节的第一个例子 (其中 $\lambda = \alpha$) 得到, 并经过变换时间尺度得到非零影响.

于是有

$$\int_{-\infty}^{\infty} \left(1-y^2\right) x\rho^{\infty}(y)\,\mathrm{d}y = 0,$$

其中 $\rho^{\infty}(y)$ 为 OU 过程的不变测度, 即是一个标准单位正态分布, 从而适用于本章提出的理论.

过程 $\varphi_\xi^t(\cdot) = \varphi^t(\cdot)$ 的生成子为

$$\mathcal{L}_0 = -\alpha y\frac{\partial}{\partial y} + \alpha\frac{\partial^2}{\partial y^2}, \tag{11.7.3}$$

并且单胞问题 (Poisson 方程) 变为

$$\alpha y\frac{\partial\Phi}{\partial y} - \alpha\frac{\partial^2\Phi}{\partial y^2} = \left(1-y^2\right)x.$$

此方程的唯一中心解为

$$\Phi(y,x) = \frac{1}{2\alpha}\left(1-y^2\right)x.$$

在标准正态分布情况下, 第四阶矩和第二阶矩取值分别为 3 和 1, 从而极限方程 (11.3.6) 的系数为

$$F(x) = \int_{-\infty}^{\infty}\left(-\frac{1}{2\alpha}y^2\left(1-y^2\right)x\right)\rho^{\infty}(y)\mathrm{d}y = \frac{1}{\alpha}x$$

且

$$A^2(x) = 2\int_{-\infty}^{\infty}\left(-\frac{1}{2\alpha}y^2x(1-y^2)x\right)\rho^{\infty}(y)\mathrm{d}y = \frac{2}{\alpha}x^2,$$

因而均匀化随机积分方程为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{X}{\alpha} + \sqrt{\frac{2}{\alpha}}X\frac{\mathrm{d}W}{\mathrm{d}t}. \tag{11.7.4}$$

这是一个在例 6.4 中研究过的几何 Brownian 运动, 其解为

$$X(t) = X(0)\exp\left(\sqrt{\frac{2}{\alpha}}W(t)\right).$$

它不收敛于 0 和 ∞, 但其时间序列有双极限. 在 10.6 节的第一个例子中, 如果 $\lambda > \alpha(\lambda < \alpha)$ 就会产生衰减 (增长), 将本例与其比较发现, 该例是对应于 $\lambda = \alpha$ 的情形, 其中变换时间尺度以获取非平凡的动力系统, 因此, 其介于衰减和增长之间. 注意到可能在带负号的白噪声之前已经得到了这个函数, 见注 11.2.

现在用替换表达式 (11.6.1) 和 (11.6.2) 得到均匀化方程的系数. 为此, 需要研究 (11.5.1) 的解变量 $\varphi^t(y)$. 由例 6.19 中给出的计算得

$$\varphi^t(y) = \mathrm{e}^{-\alpha t}y + \sqrt{2\alpha}\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}V(s),$$

$$\begin{aligned}\varphi^t(y)^2 &= \mathrm{e}^{-2\alpha t}y^2 + \sqrt{2\alpha}y\mathrm{e}^{-\alpha t}\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}V(s)\\ &\quad + 2\alpha\left(\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}V(s)\right)^2.\end{aligned}\tag{11.7.5}$$

此外, 由 Itô 等距性得

$$\mathbb{E}\left(\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}V(s)\right)^2 = \int_0^t \mathrm{e}^{-2\alpha(t-s)}\mathrm{d}s = \frac{1}{2\alpha}\left(1-\mathrm{e}^{-2\alpha t}\right).$$

为了构造测度 $\mathbb{E}^{\mu_x}$, 取初值条件 y 为一个标准的 Gaussian 分布和一个独立的 Brownian 运动 V, (事实上, 在这个特例中测度是独立于 x 的.) 从而由初始 Gaussian 分布条件下的稳定性有

$$\int \rho^\infty(y)y^2\mathrm{d}y = 1, \quad \mathbb{E}^{\mu_x}\varphi^t(y)^2 = 1,$$

而且

$$\begin{aligned}\mathbb{E}^{\mu_x}\left(\int \rho^\infty(y)y^2\varphi^t(y)^2\mathrm{d}y\right) &= \mathrm{e}^{-2\alpha t}\int \rho^\infty(y)y^4\mathrm{d}y + 2\alpha\mathbb{E}^{\mu_x}\left(\int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}V(s)\right)^2\\ &= 3\mathrm{e}^{-2\alpha t} + 1 - \mathrm{e}^{-2\alpha t}\\ &= 1 + 2\mathrm{e}^{-2\alpha t}.\end{aligned}$$

由于 $f_0(x,y) = (1-y^2)x$, 将这些计算结果代入 (11.6.2) 中得

$$\begin{aligned}F_0(x) &= x\int_0^\infty \mathbb{E}^{\mu_x}\left((1-\varphi^t(y)^2)(1-y^2)\right)\mathrm{d}t\\ &= x\int_0^\infty 2\mathrm{e}^{-2\alpha t}dt\\ &= \frac{x}{\alpha}.\end{aligned}\tag{11.7.6}$$

类似地, 由 (11.6.1) 得

$$A_0(x) = \frac{2x^2}{\alpha}.$$

这就证明有效方程为 (11.7.4).

11.7.2 快速混沌噪声

现在考虑一个完全确定的例子, 但是当消去一个快速混沌变量时, 此例就表现出随机特性. 此时, 有必要使用结论 11.7 给出的有效扩散系数方程的替换表达式. 这个表达式使用了时间积分, 并且没有涉及不变测度上的平均 (在本例中, 该不变测度不存在关于 Lebesgue 测度的密度, 见例 4.16). 考虑方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x - x^3 + \frac{\lambda}{\varepsilon}y^2, \tag{11.7.7}$$

$$\begin{aligned}
\frac{\mathrm{d}y_1}{\mathrm{d}t} &= \frac{10}{\varepsilon^2}(y_2 - y_1), \\
\frac{\mathrm{d}y_2}{\mathrm{d}t} &= \frac{1}{\varepsilon^2}(28y_1 - y_2 - y_1y_3), \\
\frac{\mathrm{d}y_3}{\mathrm{d}t} &= \frac{1}{\varepsilon^2}\left(y_1y_2 - \frac{8}{3}y_3\right).
\end{aligned} \tag{11.7.8}$$

向量 $y = (y_1, y_2, y_3)^{\mathrm{T}}$ 为 Lorenz 方程的解, 该方程参数值使得其解具有遍历性 (见例 4.16). 在不变测度中, y_2 的均值为零, 因此, 中心条件成立. x 的方程是一个由特征时间为 ε^2 的混沌信号产生的标量常微分方程. 因为 $f_0(x, y) \propto y_2$, 其不变测度如图 4.2 所示, 并且因为只有 $f_1 = (x, y) = f_1(x)$, 近似动力系统的替代方程为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = X - X^3 + \sigma\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{11.7.9}$$

其中 σ 是一个常数. 设 $\psi^t(y) = e_2 \cdot \varphi^t(y)$, 则由 (11.6.2) 给出如下常数 σ:

$$\sigma^2 = 2\lambda^2 \int_0^\infty \frac{1}{T}\left(\lim_{T\to\infty}\int_0^T \psi^s(y)\psi^{t+s}(y)\mathrm{d}s\right)\mathrm{d}t.$$

这就是 y_2 的积分自相关函数. 通过遍历性, 希望 σ^2 的值独立于 y, 并仅由 Lorenz 方程的 SRB 测度来决定. 注意到即使由于快速过程的生成子不具有椭圆性, 使得此处提出的单胞问题不具有适定性, 但是仍然希望公式是有意义的.

下面用另外一种方式推导本结论. Brownian 运动的时间导数 Gaussian 白噪声 $\sigma\dot{W}$ 可能被认为是一个与 delta 相关的平稳过程, 其自相关函数在 $[0, \infty)$ 上积分得到 $\sigma^2/2$. 假定 y_2 具有在时间上衰减的相关函数, 并注意到其有时间尺度 ε^2, 那么 $\frac{\lambda}{\varepsilon}\psi^{s/\varepsilon^2}(y)$的自相关函数在时滞 t 上是可以计算的, 在 0 到 ∞ 上对其进行积分, 并将结果与已知的关于 Gaussian 白噪声的结论进行比较, 从而得到想要的关于 σ^2 的结论.

11.7.3 Stratonovich 修正

当用光滑过程逼近白噪声时经常引出随机积分的 Stratonovich 解释, 并且至少

是一维的. 通过简单的例子, 用多尺度分析来解释这种现象. 考虑方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} f(x)\, y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{\alpha y}{\varepsilon^2} + \sqrt{\dfrac{2\alpha}{\varepsilon^2}} \dfrac{\mathrm{d}V}{\mathrm{d}t}, \end{cases} \tag{11.7.10}$$

其中 V 是一个一维标准 Brownian 运动.

简单假设 $y(0) = 0$, 则

$$\mathbb{E}(y(t)\, y(s)) = \mathrm{e}^{-\frac{\alpha}{\varepsilon^2}|t-s|}$$

因此,

$$\lim_{\varepsilon \to 0} \mathbb{E}\left(\frac{y(t)}{\varepsilon} \frac{y(s)}{\varepsilon} \right) = \frac{2}{\alpha} \delta(t-s),$$

其暗含启发式

$$\lim_{\varepsilon \to 0} \frac{y(t)}{\varepsilon} = \sqrt{\frac{2}{\alpha}} \frac{\mathrm{d}V}{\mathrm{d}t}. \tag{11.7.11}$$

理解式 (11.7.11) 的另一种方法是解 (11.7.10) 中的 y/ε,

$$\frac{y}{\varepsilon} = \sqrt{\frac{2}{\alpha}} \frac{\mathrm{d}V}{\mathrm{d}t} - \frac{\varepsilon}{\alpha} \frac{\mathrm{d}y}{\mathrm{d}t}. \tag{11.7.12}$$

如果忽略方程右边的 $O(\varepsilon)$ 项, 则可得到 (11.7.11).

上述两个结论便得猜想极限方程的形式如下:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \sqrt{\frac{2}{\alpha}} f(X) \frac{\mathrm{d}V}{\mathrm{d}t}. \tag{11.7.13}$$

在实际应用中, 可以证明由那个启发式得出的极限方程是不正确的. 这是因为在一维情形, 无论何时用光滑过程逼近白噪声, 极限方程应当在 Stratonovich 的意义下进行解释, 而在一维情形时 Stratonovich 有下面的方程:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \sqrt{\frac{2}{\alpha}} f(X) \circ \frac{\mathrm{d}V}{\mathrm{d}t}. \tag{11.7.14}$$

现在通过本章引入的方法来推导上述方程.

单胞问题

$$-\mathcal{L}_0 \Phi(x, y) = f(x)\, y,$$

其中 $\mathcal{L}_0$ 由 (11.7.3) 给出. 易得其解为

$$\Phi(x, y) = \frac{1}{\alpha} f(x)\, y, \quad \nabla_x \Phi(x, y) = \frac{1}{\alpha} f'(x)\, y.$$

不变密度为

$$\rho^{\infty}(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right),$$

该密度在 $\mathcal{L}_0^*$ 的零空间上, 并对应于一个标准的 Gaussian 随机变量 $\mathcal{N}(0,1)$.

由方程 (11.3.2) 得

$$F(x) = \int_{\mathbb{R}} \frac{1}{\alpha} f'(x) f(x) y^2 \rho^{\infty}(y) \mathrm{d}y = \frac{1}{\alpha} f'(x) f(x).$$

同样由 (11.3.3) 得

$$A(x)^2 = \int_{\mathbb{R}} \frac{2}{\alpha} f(x)^2 y^2 \rho^{\infty}(y) \mathrm{d}y = \frac{2}{\alpha} f(x)^2.$$

因此, 极限方程是 Itô 随机微分方程

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{1}{\alpha} f'(X) f(X) + \sqrt{\frac{2}{\alpha}} f(X) \frac{\mathrm{d}V}{\mathrm{d}t}.$$

由注 6.2 知这是 (11.7.14) 的 Itô 形式, 因此, 结论成立.

11.7.4 Stokes 定理

前面的例子可视为是在描述一个无质量粒子的运动, 该运动在速度场中 x 的位置与 $f(x)y$ 对称, 其中 y 是一个 OU 过程. 如果粒子的质量为 m, 则通常研究如下广义方程:

$$\begin{cases} m\dfrac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \dfrac{1}{\varepsilon} f(x) y - \dfrac{\mathrm{d}x}{\mathrm{d}t}, & (11.7.15\text{a}) \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{\alpha y}{\varepsilon^2} + \sqrt{\dfrac{2\alpha}{\varepsilon^2}} \dfrac{\mathrm{d}V}{\mathrm{d}t}. & (11.7.15\text{b}) \end{cases}$$

(注意设 $m = 0$ 可得到前面的例子.) 方程 (11.7.15a) 是 Stokes 定理, 表明作用在粒子上的力与拉力成比例,

$$\frac{1}{\varepsilon} f(x) y - \frac{\mathrm{d}x}{\mathrm{d}t},$$

上式相当于流体场与粒子场之间的区别. 正如前面的例子一样, y 是一个起伏的 OU 过程. 为了简便起见, 考虑单位质量 $m = 1$ 的情形.

由前面章节中启发式的结论自然想到如下形式的极限方程:

$$\frac{\mathrm{d}^2 X}{\mathrm{d}t^2} = \sqrt{\frac{2}{\alpha}} f(X) \frac{\mathrm{d}V}{\mathrm{d}t} - \frac{\mathrm{d}X}{\mathrm{d}t}. \tag{11.7.16}$$

与前面的应用相比, 这个猜想的极限方程是正确的, 因为此处 x 更加光滑, 并且 Itô 与 Stratonovich 积分一致, 这里没有对 Stratonovich 积分进行 Itô 修正 (需

将 (11.7.16) 记为一阶系统来理解此结论, 见练习第 2 题 (a)). 使用本章引入的多尺度方法证明这个结论.

首先将 (11.7.16) 记为如下一阶系统:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = r, \\ \dfrac{\mathrm{d}r}{\mathrm{d}t} = -r + \dfrac{1}{\varepsilon} f(x) y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{1}{\varepsilon^2}\alpha y + \dfrac{1}{\varepsilon}\sqrt{2\alpha}\dfrac{\mathrm{d}V}{\mathrm{d}t}. \end{cases}$$

其中 (x, r) 是慢变量 (x 在 (11.2.1) 中), 并且 y 是快变量 (y 在 (11.2.1) 中). 由下式给出单胞问题:

$$\mathcal{L}_0 \Phi(x, r, y) = -f_0(x, r, y) = \begin{pmatrix} 0 \\ -f(x) y \end{pmatrix},$$

其中 $\mathcal{L}_0$ 由 (11.7.3) 得到. 上述方程的解为

$$\Phi(x, r, y) = \begin{pmatrix} 0 \\ \dfrac{1}{\alpha} f(x) y \end{pmatrix}, \quad \nabla_{(x,r)} \Phi(x, y) = \begin{pmatrix} 0 & 0 \\ \dfrac{1}{\alpha} f'(x) y & 0 \end{pmatrix}.$$

注意到 f_0 在 $\nabla_{(x,r)}\Phi$ 的零空间上, 因此, 由 (11.3.2) 得到

$$F(X, R) = F_1(X, R) = \begin{pmatrix} R \\ -R \end{pmatrix}. \tag{11.7.17}$$

由 (11.3.3) 有

$$A(X, R) A(X, R)^{\mathrm{T}} = \int_{\mathbb{R}} 2 \begin{pmatrix} 0 & 0 \\ 0 & \dfrac{1}{\alpha} f(X)^2 y^2 \end{pmatrix} \rho^{\infty}(y) \mathrm{d}y.$$

回想 $\rho^{\infty}(y)$ 是一个 Gaussian 随机变量 $\mathcal{N}(0,1)$ 的密度, 积分得

$$A(X, R) A(X, R)^{\mathrm{T}} = \begin{pmatrix} 0 & 0 \\ 0 & \dfrac{2}{\alpha} f(X)^2 \end{pmatrix}.$$

所以, 通常选择 $A(x)$ 为

$$A(X, R) = \begin{pmatrix} 0 \\ \sqrt{\dfrac{2}{\alpha}} f(X) \end{pmatrix}.$$

则根据 (11.7.17) 和 (11.7.18) 得到下面的极限方程:

$$\begin{cases} \dfrac{\mathrm{d}X}{\mathrm{d}t} = R, \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = -R + \sqrt{\dfrac{2}{\alpha}} f(X) \dfrac{\mathrm{d}W}{\mathrm{d}t}, \end{cases}$$

消除 R 可见其与猜想的极限方程 (11.7.16) 一致.

11.7.5 Green-Kubo 公式

在前面应用中讨论的是具有明显质量的粒子的运动方程, 其服从 Stokes 拉力. 这里研究同样的运动方程, 但是其速度场是稳定的. 同样假设粒子服从分子扩散, 则运动方程为

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = f(x) - \frac{\mathrm{d}x}{\mathrm{d}t} + \sigma \frac{\mathrm{d}U}{\mathrm{d}t}, \tag{11.7.18}$$

其中 U 是标准的 Brownian 运动. 假设 $f(x)$ 是均值为零的周期函数, 将在大长度和时间尺度上研究粒子 x 的有效扩散行为. 在恰当的大长度和时间尺度上, 证明粒子在做一个有效的 Brownian 运动, 并计算其扩散系数.

为此, 通过设 $x \to x/\varepsilon$ 和 $t \to t/\varepsilon^2$ 来重新获取运动方程,

$$\varepsilon^2 \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \frac{1}{\varepsilon} f\left(\frac{x}{\varepsilon}\right) - \frac{\mathrm{d}x}{\mathrm{d}t} + \sigma \frac{\mathrm{d}U}{\mathrm{d}t}.$$

引入变量 $y = \varepsilon \mathrm{d}x/\mathrm{d}t$ 和 $z = x/\varepsilon$, 得到如下系统:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{1}{\varepsilon^2} y + \dfrac{1}{\varepsilon^2} f(z) + \dfrac{\sigma}{\varepsilon} \dfrac{\mathrm{d}W}{\mathrm{d}t}, \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2} y. \end{cases}$$

过程 (y,z) 具有遍历性, 其特征时间尺度为 ε^2, y 在方程 (11.2.1) 中仍然起着 y 的作用; x 在方程 (11.2.1) 中仍然起着 x 的作用. 算子 $\mathcal{L}_0$ 是过程 (y,z) 的生成子, 而且

$$f_1(x,y,z) = 0, \quad f_0(x,y,z) = y.$$

因此, 由于 (y,z) 的进化过程是独立于 x 和 $\Phi(x,y,z)$ 的, 所以单胞问题的解也是独立于 x 的, 从而由 (11.3.2) 得 $F(x) = 0$. 考虑有效扩散系数时发现, 由于 $\alpha(x,y) = A(x,y) = 0$, 由 (11.3.3) 得出 $A(x)^2 = A_0(x)$. 现在定义 $\psi^t(y,z)$ 为 $\varphi^t(y,z)$ 的一部分, 表示 $\varphi^t(y,z)$ 投影到 y 坐标上的投影值. 由结论 11.7 有

$$A_0(x) = 2\int_0^\infty \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \psi^s(y)\,\psi^{s+t}(y)\,\mathrm{d}s \right) \mathrm{d}t.$$

表达式

$$C(t)=\lim_{T\to\infty}\frac{1}{T}\int_0^T\psi^s(y)\,\psi^{s+t}(y)\,\mathrm{d}s$$

是速度自相关函数, 则有效方程

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\sqrt{2D}\frac{\mathrm{d}W}{\mathrm{d}t}$$

是一个具有如下扩散系数的 Brownian 运动:

$$D=\int_0^\infty C(t)\mathrm{d}t,$$

则由积分速度自相关性可得到有效扩散系数. 这是 **Green-Kubo 公式**的一个例子.

11.7.6 非 Itô, Stratonovich 情形

再次对小质量 $m=\tau_0\varepsilon^2$ 的粒子使用 Stokes 定理 (11.7.15a), 其中 $\tau_0=O(1)$, 并且不计分子扩散. 如果还假设下面流体的速度场的形式为 $\frac{1}{\varepsilon}f(x)\eta$, 其中 η 为随机微分方程的解, 则可得

$$\begin{cases}\tau_0\varepsilon^2\dfrac{\mathrm{d}^2x}{\mathrm{d}t^2}=-\dfrac{\mathrm{d}x}{\mathrm{d}t}+\dfrac{1}{\varepsilon}f(x)\,\eta, & (11.7.19a)\\[2ex] \dfrac{\mathrm{d}\eta}{\mathrm{d}t}=\dfrac{1}{\varepsilon^2}g(\eta)+\dfrac{1}{\varepsilon}\sqrt{2\sigma(\eta)}\dfrac{\mathrm{d}W}{\mathrm{d}t}. & (11.7.19b)\end{cases}$$

用 Itô 公式来解释方程 (11.7.19b). 假设 $g(\eta)$, $\sigma(\eta)$ 使 (11.7.19b) 的 Fokker-Planck 方程存在一个唯一稳态解, 这样 η 就具有遍历性.

记 (11.7.19) 为一阶系统,

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=\dfrac{1}{\varepsilon\sqrt{\tau_0}}v,\\[2ex] \dfrac{\mathrm{d}v}{\mathrm{d}t}=\dfrac{f(x)\,\eta}{\varepsilon^2\sqrt{\tau_0}}-\dfrac{v}{\tau_0\varepsilon^2},\\[2ex] \dfrac{\mathrm{d}\eta}{\mathrm{d}t}=\dfrac{g(\eta)}{\varepsilon^2}+\dfrac{\sqrt{2\sigma(\eta)}}{\varepsilon}\dfrac{\mathrm{d}W}{\mathrm{d}t}.\end{cases} \tag{11.7.20}$$

方程 (11.7.20) 具有 (11.2.1) 的形式, 假设快速过程 (v,η) 具有遍历性, 则其适用于本章结论. 为了计算有效系数, 需要解如下静态 Fokker-Planck 方程:

$$\mathcal{L}_0^*\rho(x,y,\eta)=0$$

以及单胞问题

$$-\mathcal{L}_0h=\frac{v}{\sqrt{\tau_0}}, \tag{11.7.21}$$

其中

$$\mathcal{L}_0 = g(\eta)\frac{\partial}{\partial \eta} + \sigma(\eta)\frac{\partial^2}{\partial \eta^2} + \left(\frac{f(x)\eta}{\sqrt{\tau_0}} - \frac{v}{\tau_0}\right)\frac{\partial}{\partial v}.$$

简化方程 (11.7.21) 找到其具有如下形式的解:

$$h(x,v,\eta) = \left(\sqrt{\tau_0}v + f(x)\hat{h}(\eta)\right). \tag{11.7.22}$$

在单胞问题中替换式 (11.7.22), 并进行代数运算得方程

$$-\mathcal{L}_\eta \hat{h} = \eta.$$

其中 $\mathcal{L}_\eta$ 定义为 η 的生成子. 假设 $\eta(t)$ 的唯一不变测度关于 Lebesgue 测度的密度为 $\rho_\eta(\eta)$, 保证 $\hat{h}$ 的 Poisson 方程具有适定性的中心条件是

$$\int_{\mathbb{R}} \eta \rho_\eta(\eta)\,\mathrm{d}\eta = 0,$$

假设其是成立的. 均匀随机微分方程为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) + \sqrt{D(X)}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{11.7.23}$$

其中

$$F(x) := \int_{\mathbb{R}^2}\left(\frac{v}{\sqrt{\tau_0}}\hat{h}(\eta)f'(x)\right)\rho(x,v,\eta)\,\mathrm{d}v\mathrm{d}\eta$$

且

$$D(x) := 2\int_{\mathbb{R}^2}\left(v^2 + \frac{v}{\sqrt{\tau_0}}\hat{h}(\eta)f(x)\right)\rho(x,v,\eta)\,\mathrm{d}v\mathrm{d}\eta.$$

当 $\eta(t)$ 是一个 Ornstein-Uhlenbeck 过程时,

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\frac{\alpha}{\varepsilon^2}\eta + \sqrt{\frac{2\lambda}{\varepsilon^2}}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{11.7.24}$$

可以显式地计算出均匀系数 $D(X)$ 和 $B(X)$. 有效随机微分方程为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\lambda}{\alpha^2(1+\tau_0\alpha)}f(X)f'(X) + \sqrt{\frac{2\lambda}{\alpha^2}}f(X)\frac{\mathrm{d}W}{\mathrm{d}t}. \tag{11.7.25}$$

注意到当 $\tau_0 \to \infty$ 时, 又得到了 11.7.4 小节中出现过的 Itô 随机积分; 反之, 当 $\tau_0 \to 0$ 时, 得到 11.7.3 小节中出现过的 Stratonovich 随机积分的 Itô 解释. 对于 $\tau_0 \in (0,\infty)$, 极限方程既不是 Itô 形式, 也不是 Stratonovich 形式. 事实上, 方程 (11.7.25) 可记为如下形式:

$$X(t) = x_0 + \int_0^t \frac{2\lambda}{\alpha^2}f(X)\,\hat{\circ}\mathrm{d}W(t),$$

此处通过 Riemann 求和定义的随机积分依赖于 τ_0 的值. 恢复这一个有趣的极限的事实是与质量的尺度 $O\left(\varepsilon^2\right)$ 密切联系在一起的. 这种尺度换算确保遍历性过程 η 的时间尺度以及粒子的松弛时间是相同的. 这些时间尺度间的共振给出想要的效果.

11.7.7 Lévy 面积修正[①]

在 11.7.3 小节中看到在一维情形下对白噪声进行光滑估计时产生了 Stratonovich 随机积分. 通常这是不可能的, 但是在多维情形下会产生这种情况, 即极限中可能出现一个额外的漂移项. 这个额外的漂移对极限过程中的 Lévy 面积特性有贡献 (见 11.8 节).

考虑快 – 慢系统

$$\dot{x}_1 = \frac{1}{\varepsilon} y_1, \tag{11.7.26a}$$

$$\dot{x}_2 = \frac{1}{\varepsilon} y_2, \tag{11.7.26b}$$

$$\dot{x}_3 = \frac{1}{\varepsilon}\left(x_1 y_2 - x_2 y_1\right), \tag{11.7.26c}$$

$$\dot{y}_1 = -\frac{1}{\varepsilon^2} y_1 - \alpha \frac{1}{\varepsilon^2} y_2 + \frac{1}{\varepsilon}\dot{W}_1, \tag{11.7.26d}$$

$$\dot{y}_2 = -\frac{1}{\varepsilon^2} y_2 + \alpha \frac{1}{\varepsilon^2} y_1 + \frac{1}{\varepsilon}\dot{W}_2, \tag{11.7.26e}$$

其中 $\alpha > 0$, W_1, W_2 是标准独立 Brownian 运动.

注意方程 (11.7.26d) 和 (11.7.26e) 可记为如下形式:

$$\dot{y} = -\frac{1}{\varepsilon^2} y + \frac{1}{\varepsilon^2}\alpha J_y + \frac{1}{\varepsilon}\dot{W},$$

其中 $y = (y_1, y_2)$, $W = (W_1, W_2)$, 并且 J 是反对称 (辛) 矩阵,

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0. \end{pmatrix}.$$

运用启发式

$$y \approx \varepsilon\left(I - \alpha J\right)^{-1}\frac{\mathrm{d}W}{\mathrm{d}t}$$

得到极限方程

$$\dot{x}_1 = \frac{1}{1+\alpha^2}\left(\dot{W}_1 - \alpha\dot{W}_2\right), \tag{11.7.27a}$$

$$\dot{x}_2 = \frac{1}{1+\alpha^2}\left(\dot{W}_2 + \alpha\dot{W}_1\right), \tag{11.7.27b}$$

① 本节是与 M.Hairer 一起编写的.

$$\dot{x}_3 = \frac{1}{1+\alpha^2}\left((\alpha x_1 - x_2)\dot{W}_1 + (\alpha x_2 + x_1)\dot{W}_2\right). \tag{11.7.27c}$$

由 11.7.3 小节和 11.7.6 小节可知, 当白噪声的典型光滑逼近引起 Stratonovich 随机积分时, 必须注意猜测这样一个极限. 然而, 在上述情形下, 由于 Itô 和 Stratonovich 一致使得这种情况不会发生. 尽管如此, 那个猜测出来的极限方程是错误的.

使用本章叙述的多尺度方法时会得到如下正确的均匀化系统:

$$\dot{x}_1 = \frac{1}{1+\alpha^2}\left(\dot{W}_1 - \alpha\dot{W}_2\right), \tag{11.7.28a}$$

$$\dot{x}_2 = \frac{1}{1+\alpha^2}\left(\dot{W}_2 + \alpha\dot{W}_1\right), \tag{11.7.28b}$$

$$\dot{x}_3 = \frac{1}{1+\alpha^2}\left((\alpha x_1 - x_2)\dot{W}_1 + (\alpha x_2 + x_1)\dot{W}_2\right) + \frac{\alpha}{1+\alpha^2}. \tag{11.7.28c}$$

注意到在方程 (11.7.28c) 中出现了附加常数漂移, 它是快速过程 y 方程中的反对称部分, 快速过程 y 直接导致均匀化方程中附加漂移的出现. 特别地, 当 $\alpha = 0$ 时, 均匀化方程变为

$$\dot{x}_1 = \dot{W}_1,$$

$$\dot{x}_2 = \dot{W}_2,$$

$$\dot{x}_3 = -x_2\dot{W}_1 + x_1\dot{W}_2,$$

该方程与开始 (一般来说, 是不正确的) 时猜测的极限 (11.7.27) 一致.

11.8 讨论和参考

由 Papanicolaou 及其合作者撰写的论文中讨论了本章使用的扰动展开方法以及一些更常用的相关方法, 见文献 [240 ~ 242, 244], 这些方法都是建立在 Khasminkii 研究的基础上, 见文献 [165, 166], 在文献 [155, 154, 31, 205, 240, 242, 244] 中可以找到进一步的研究资料. 使用一般的方法分析 $\mathcal{Y} = \mathbb{T}^d$ 这种简单情形, 通过设置恰当的函数, 也可以将这种方法推广到 $\mathcal{Y} = \mathbb{R}^d$ 的例子上, 见文献 [249 ~ 251].

本章列出的基本扰动展开被证明是严格正确的, 并且当 $\varepsilon \to 0$ 时 x 弱收敛于 X, 见文献 [181] 和第 18 章; 在文献 [241] 中明确地提出了构成这种方法的扰动展开, 也可见文献 [117,ch.6,291,321]; 对快速过程的 Fokker-Planck 算子使用特征函数展开也可以分析类似的问题, 见文献 [27,ch.8]; 在物理文献中通常使用投影算子方法来消去快速变量, 见文献 [117,ch.6] 及其参考文献.

当初始系统是一个常微分方程时, 研究如何推导有效随机模型是一类问题, 可在文献 [242] 中的一些概论中找到这方面的研究. 11.7.2 小节中出现的具体例子依

赖于 Lorenz 方程的遍历性, 这些在 [318, 319] 中得到证实. 文献 [322] 重点介绍了在数值上用积分自相关函数计算有效扩散系数; 用另一种方法计算有效扩散系数见 [125]. 为了便于分析, Beck 使用了绕积结构, 并在离散时间上实现了被描述的程序, 见 [31], 后来, 在一些例子中证明这种思想是正确的. 在文献 [322] 和文献 [125] 中也采用了绕积设置. [225] 中证明了由快速混合系统所导致的常微分方程的严格极限理论, 其中使用了 [224] 中发展的动力系统的大偏差原理. 在论文 [208] 中, 采用了这样一种思想: 快速混沌特性能够用慢变量来引出噪声, 而这一种思想源自于文献 [204] 中的 Burgers. 有关 Burgers 洗浴模型进一步的数值实验见文献 [209].

对于微分方程的相关工作可在文献 [124] 中找到, 并且关于连续时间的类似思想也在文献 [154, 155] 中进行了讨论; 然而, 通过使用文献 [65] 中的 Mori-Zwanzig 形式, 找到 Fokker-Planck 方程投影到 $\mathcal{X}$ 空间上的严格解, 再在结果中对 ε 进行幂级数展开, 而不是直接针对 ε 的幂进行系统展开.

11.7.5 小节中, 通过速度自相关函数积分推导出有效扩散系数的计算公式, 得到了 Green-Kubo 公式. 这样就通过自相关函数的时间积分计算出一个传递系数. 在许多统计数学的书籍中都对 Green-Kubo 公式以及其他传递系数的计算进行了研究, 见文献 [28,ch.11,269].

相对变化慢的海洋变量, 大气变化得很快, 在文献 [202, 205] 中, Majda 等运用多尺度方法对这种气候模式进行了分析. 大气科学中关于多尺度方法更进一步应用见文献 [206, 207], 也可见文献 [78]. Stokes 定理 (方程 (11.7.15a)) 是一个流体场中惯性粒子运动的表象模型, 见文献 [217]. 在文献 [288, 289] 中对形式如 (11.7.15) 的模型进行了理论延伸, 并在文献 [254] 中有进一步的分析, 在此模型中, 粒子浸入的流体速度场被看成是一个 Gaussian Markovian 随机场. 对无源示踪器中类似的 Gaussian 模型的研究见 [55, 56].

事实上, 在一维情形下对白噪声进行光滑逼近 (在极限处消除了正则化) 得到的 Stratonovich 积分, 通常称其为 Wong-Zakai 定理, 见文献 [332]. 不管是用 Itô 还是用 Stratonovich 来解释随机积分都称之为 Itô 与 Stratonovich 问题. 正如 11.7.6 小节中考虑的例子, 在出现不止一个快时间尺度的情形下, 正确地解释极限随机积分方程中的随机积分依赖于求极限时的阶数, 见文献 [109, 280]. 如 11.7.6 小节所示, 可以既不用 Itô 也不用 Stratonovich 来解释许多极限随机积分方程中随机积分的例子, 见文献 [129, 180, 255]. [124] 考察了快速过程是一个离散决定性混沌映射的问题. 在这种情况下, 考虑一种有趣的设置, 即 Stokes 定理 (11.7.15), 此时, 质量是很小的,

$$\varepsilon^a \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \frac{1}{\varepsilon} f(x) y - \frac{\mathrm{d}x}{\mathrm{d}t} + \sigma \frac{\mathrm{d}U}{\mathrm{d}t},$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{\alpha y}{\varepsilon^2} + \sqrt{\frac{2\alpha}{\varepsilon^2}} \frac{\mathrm{d}V}{\mathrm{d}t}.$$

在第一个方程中设 $\varepsilon = 0$, 并且对 y/ε 使用白噪声逼近时, 猜测 x 的极限 X 满足一阶随机微分方程, 问题就变成了对随机积分的解释. 在 [180] 中用多尺度展开推导当 $a = 1, 2, 3$ 时满足 x 的极限方程. 当 $a = 1$ 时, 极限是一个 Itô 方程; 当 $a = 3$ 时, 极限是一个 Stratonovich 方程; 当 $a = 2$ 时, 极限介于上述两种情形之间.

在更高维对白噪声进行光滑逼近 (通常依赖于正则化的类型) 时, 除了 Stratonovich 随机积分之外, 还会产生一个与扩散矩阵行向量间通信有关的附加漂移, 见文献 [151]. 理解 11.7.7 小节中所示基于粗糙路径的例子, 其严格理论框架见文献 [198].

本章中已经考虑了形如 (11.2.1) 的方程, 其中 U 和 V 是独立的 Brownian 运动, 通常由两个过程中噪声的相互关联得到其应用. 将在第 13 章中讨论这种情形. 已讨论的线性方程的结构通常足以包含向后 Kolmogorov 方程的形式, 当 U 和 V 相关联 (实际上是相同的) 时, 由 (11.2.1) 得到向后 Kolmogorov 方程. 在本章的推导过程中, 主要的变化是生成子 $\mathcal{L}_1$ 包含由噪声的相互关系得到的附加项, 见练习第 1 题和第 5 题.

当写出完全系统的向后 Kolmogorov 方程 (11.2.2) 时, 假设初始条件仅依赖于慢变量 x. 这个假设其实是没有必要的, 它仅仅是简化了分析. 如果初始条件是 x 和 y 的函数, 则会出现一个需要重新解答的初始 (或者是边界) 层. 通过在双尺度展开式 (该展开式在时间上呈指数衰减) 上增加恰当的项可以实现这个想法. 文献 [336] 和文献 [167] 中分别对连续时间的 Markov 链以及随机微分方程实现了这种想法. 此时对在快速过程不变测度下的随机微分方程的初始条件平均化, 可以得到均匀化随机微分方程的初始条件.

本章已经研究了有限维随机系统的均匀化, 对于无限维随机系统以及随机偏微分方程也可证明类似的结论. 用本章的方法来求解随机 Burgers 方程见文献 [40].

结论 11.1 中表达式的应用见文献 [241]. 结论 11.7 和结论 11.6 中关于有效漂移和扩散的表达式可用于设计粗略的时间步长算法, 见文献 [322]. 通常, 在随机微分方程 (11.2.1) 中出现两个广泛分离的特征时间尺度时, 使得求其数值解很困难. 现在已经研究出针对这种问题的有效数值解的解法, 在文献 [116] 中描述了 Hamiltonian 系统平均背景下的这个话题, 在更一般的情况下, 还会遇到这个话题, 见文献 [93]. 这些方法中的绝大多数都用到如下事实, 即当 ε 足够小时, (11.2.1a) 的解可由均匀化方程 (11.3.6) 的解逼近. 通过形如 (11.6.3) 或 (11.6.1) 的公式以及在短时间间隔上的积分方程 (11.2.1b) 都可以计算均匀化系数, 见文献 [81, 84, 122, 123]. [162] 列出一个动力系统 (可能是随机的) 的数值程序可以计算出变量的一个子集; 这种没有使用明显尺度分离的方法应用于解决一系列不同的问题, 见文献 [30, 149, 163, 164, 190, 278, 334]. [83] 中概括了多尺度问题的数值解法. 估计

多尺度随机微分方程中的参数见文献 [258]. 降低随机系统维数见文献 [59, 148, 273].

11.9 练　　习

1. 假定 f_0 满足中心条件, 并且 U 和 V 是独立的 Brownian 运动, 找出下面随机微分方程的均匀化方程:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} f_0(x,y) + f_1(x,y) + \alpha_0(x,y)\dfrac{\mathrm{d}U}{\mathrm{d}t} + \alpha_1(x,y)\dfrac{\mathrm{d}V}{\mathrm{d}t}, \quad x(0) = x_0, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2} g(x,y) + \dfrac{1}{\varepsilon} g_1(x,y) + \dfrac{1}{\varepsilon}\beta(x,y)\dfrac{\mathrm{d}V}{\mathrm{d}t}, \quad y(0) = y_0. \end{cases}$$

2. (a) 用 $\mathcal{Y}$ 表示 $\mathbb{T}^d$ 或 $\mathbb{R}^d$. 由下面方程给出的过程 $y \in \mathcal{Y}$ 的生成子 $\mathcal{L}$ 是什么:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = g(y) + \frac{\mathrm{d}V}{\mathrm{d}t}.$$

当 $g(y) = -\nabla\Psi(y)$ 时, 在 $\mathcal{L}^*$ 的零空间上找到一个函数.

(b) 当 $g = -\nabla\Psi(y)$ 时, 找出下面系统:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} f(x,y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2} g(y) + \dfrac{1}{\varepsilon}\dfrac{\mathrm{d}V}{\mathrm{d}t} \end{cases}$$

产生的均匀化随机微分方程.

(c) 定义单胞问题, 给出恰当的条件, 使得当 $\mathcal{Y} = \mathbb{T}^d$ 时解是唯一的. 明确在上面的推导中 f 所需要的假设.

3. 使用 Itô 公式推导随机微分方程 (11.7.4) 的解. 将这个随机微分方程转化为 Stratonovich 形式, 你观察到了什么?

4. (a) 设 $\mathcal{Y}$ 表示为 $\mathbb{T}^d$ 或 $\mathbb{R}^d$. 写出下面方程给出的过程 $y \in \mathcal{Y}$ 的生成子 $\mathcal{L}_0$:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = g(y) + \frac{\mathrm{d}V}{\mathrm{d}t}.$$

在 g 是无散的情况下, 在 $\mathcal{L}_0^*$ 的零空间上找到一个函数.

(b) 找出下面系统:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x,y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon} g(y) + \dfrac{1}{\sqrt{\varepsilon}}\dfrac{\mathrm{d}V}{\mathrm{d}t} \end{cases}$$

产生的平均随机微分方程, 其中 g 是无散的.

(c) 找出下面系统:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon} f(x,y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2} g(y) + \dfrac{1}{\varepsilon}\dfrac{\mathrm{d}V}{\mathrm{d}t} \end{cases}$$

产生的均匀化随机微分方程, 其中 g 是无散的.

(d) 定义单胞问题, 给出恰当的条件, 使得当 $\mathcal{Y} = \mathbb{T}^d$ 时解是唯一的. 明确在上面的推导中 f 所需要的假设.

5. 考虑运动方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x) + \sigma\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中 $f(x)$ 是无散的且具有周期性, 均值为零. 弄明白 x 如何在大长度和时间尺度上起作用是很有趣的. 为此, 通过设 $x \to x/\varepsilon$ 以及 $t \to t/\varepsilon^2$, 并引入 $y \to x/\varepsilon$ 来重新标定运动方程. 写出一对 x 和 y 的随机微分方程, 使用第 1 题中的方法确保消去 y 而得到 x 的有效方程.

6. 在任意维情况下给出 11.7.6 小节中的分析, 极限方程是否具有一维情形下相同的结构?

7. 当 (11.7.24) 给出 $\eta(t)$ 时, 由 (11.7.23) 推导方程 (11.7.25).

8. (Kramers Smoluchowski 极限) 考虑 Langevin 方程

$$\varepsilon^2\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = b(x) - \frac{\mathrm{d}x}{\mathrm{d}t} + \sqrt{2\sigma}\frac{\mathrm{d}W}{\mathrm{d}t} \tag{11.9.1}$$

其中假定粒子质量很小, $m = \varepsilon^2$.

(a) 通过引入变量 $y = \varepsilon\dot{x}$, 写出 (11.9.1) 的一阶系统;

(b) 使用多尺度分析证明当 $\varepsilon \ll 1$ 时, 使用如下 Smoluchowski 方程:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = b(X) + \sqrt{2\sigma}\frac{\mathrm{d}W}{\mathrm{d}t}$$

的解能很好地逼近 (11.9.1) 的解.

(c) 计算 Smoluchowski 方程的一阶校正值.

9. 写出方程 (11.7.16) 一阶系统, 并且证明此方程的 Itô 形式和 Stratonovich 形式是一致的.

第 12 章　椭圆型偏微分方程的均匀化

12.1　引　　言

本章使用多尺度展开来研究二阶散度型一致椭圆型偏微分方程的均匀化问题. 从形式上来讲, 用于推导均匀化方程的计算方法与前面章节中研究随机微分方程均匀化时用到的方法非常相似, 主要的区别在于所研究的线性方程中没有时间依赖性.

在 12.2 节中给出本章研究的边界值问题. 12.3 节中包含简化 (均匀) 方程, 并在 12.4 节中给出其推导过程. 在 12.5 节中研究了简化方程的结构, 表明其继承了原始方程的椭圆性. 在 12.6 节中描述了该理论的两种应用, 并且都具有显式解, 一个是一维的例子, 一个是二维分层介质.

12.2　完 全 方 程

研究如下散度型一致椭圆偏微分方程, 其具有 Dirichlet 边界条件:

$$-\nabla\cdot\left(A^{\varepsilon}\nabla u^{\varepsilon}\right)=f,\quad x\in\Omega, \tag{12.2.1a}$$

$$u^{\varepsilon}=0,\quad x\in\partial\Omega, \tag{12.2.1b}$$

其中 $u^{\varepsilon}=u^{\varepsilon}(x)$ 是一个有待确定的未知标量场, $A^{\varepsilon}=A(x/\varepsilon)$ 是一个已给矩阵, $f=f(x)$ 是一个已给的标量场. 不同于前面 4 章, 这里不存在两个不同的显式变量 x 和 y. 类似于前面章节, 设 $y=x/\varepsilon$, 目的是要推导一个均匀化方程, 其中当 $\varepsilon\to0$ 时可以消去 y, 而且要研究均匀化方程的不同特性.

令 $\Omega\subset\mathbb{R}^d$ 是开集且具有光滑边界. 假设矩阵值函数 $A(y)$ 是光滑的、周期为 1 且一致正定. 这个假设暗含方程 (12.2.1a) 左边出现的微分算子是一致椭圆的 (见第 7 章), 而且令函数 $f(x)$ 是光滑的, 并且独立于 ε. 简言之, 有如下假设:

$$f\in C^{\infty}\left(\mathbb{R}^d,\mathbb{R}\right), \tag{12.2.2a}$$

$$A\in C^{\infty}_{\text{per}}\left(\mathbb{T}^d,\mathbb{R}^{d\times d}\right), \tag{12.2.2b}$$

$$\exists\alpha>0:\langle\xi,A(y)\xi\rangle\geqslant\alpha\left|\xi\right|^2,\forall y\in\mathbb{T}^d,\forall\xi\in\mathbb{R}^d. \tag{12.2.2c}$$

注意到对于独立于 ε 的 β 和 α, A 的假设暗含了 $A^{\varepsilon} \in M(\alpha, \beta, \Omega)$. 这种正则性假定不仅是必要的, 更是严格的, 这种假设是为了进行随后的计算. 作最小正则假设是一个很重要的问题, 然而在许多应用中, 人们希望当从一种材料穿越到另一种材料时, 系数 $A(y)$ 已经发生跳跃. 第 19 章中对均匀化理论的证明将减弱此处所作的正则性假设.

令 $\mathcal{A}_0 = -\nabla_y \cdot (A\nabla_y)$ 在单位环面上具有周期边界条件且 $A = A(y)$. 如例 7.12 中所研究的那样, 这个算子在下面的讨论中将起到关键作用. 该例中还表明算子具有一维零空间, 并在 y 上含有常数, 而且使用 Fredholm 性质表明 Poisson 方程

$$\mathcal{A}_0 v = h, \quad v\text{是以 1 为周期的} \tag{12.2.3}$$

有解的充要条件为

$$\int_{\mathbb{T}^d} h(y)\,\mathrm{d}y = 0, \tag{12.2.4}$$

其解的唯一性依赖于一个附加常数. 在所有 (12.2.3) 的解 (均满足可解条件) 中选取唯一解, 其在 $\mathbb{T}^d$ 上的积分为零,

$$\mathcal{A}_0 v = h, \quad v\text{是以 1 为周期的}, \quad \int_{\mathbb{T}^d} v(y)\,\mathrm{d}y = 0.$$

形如 (12.2.3) 的方程将在其后的讨论中扮演重要的角色.

12.3 简 化 方 程

用如下公式定义**有效扩散张量**:

$$\bar{A} = \int_{\mathbb{T}^d} \left(A(y) + A(y)\nabla\chi(y)^{\mathrm{T}}\right)\mathrm{d}y, \tag{12.3.1}$$

其中向量场 $\chi : \mathbb{T}^d \to \mathbb{R}^d$ 满足**单胞问题**

$$-\nabla_y \cdot \left(\nabla_y \chi A^{\mathrm{T}}\right) = \nabla_y \cdot A^{\mathrm{T}}, \quad \chi\text{是以 1 为周期的}. \tag{12.3.2}$$

结论 12.1 对于 $0 < \varepsilon \ll 1$, 方程 (12.2.1) 的解 u^{ε} 近似地由如下均匀化方程的解 u 得到:

$$-\nabla \cdot \left(\bar{A}\nabla u\right) = f, \quad x \in \Omega, \tag{12.3.3a}$$

$$u = 0, \quad x \in \partial\Omega. \tag{12.3.3b}$$

注意: 确定场 χ 取决于一个常向量. 然而, 由于在均匀化方程中出现的均匀化矩阵 $\bar{A}$ 的计算公式中仅涉及 $\nabla_y \chi$, 因此, 这个常数值是不相关的. 但是, 就定义而言, 对 χ 进行规范化可得唯一解

$$\int_{\mathbb{T}^d} \chi(y)\,\mathrm{d}y = 0. \tag{12.3.4}$$

有时将单胞问题用其替换形式表示是有用的, 该替换形式可通过对 χ 的每个元素写一个方程而得到

$$\mathcal{A}_0\chi_\ell = \nabla_y \cdot a_\ell, \quad \ell = 1, \cdots, d, \tag{12.3.5}$$

其中 $a_\ell = Ae_\ell(\ell = 1, \cdots, d)$, 并且 $\{e_\ell\}_{\ell=1}^d$ 是 $\mathbb{R}^d$ 上的标准基. 因此, a_ℓ 为矩阵 A 的第 ℓ 列.

注 12.2 由于 Hessian 矩阵 $\nabla_x\nabla_x u$ 是对称的, 则对 (12.3.1) 使用性质 (2.2.2) 同样可以得到下面关于 $\bar{A}$ 的表达式:

$$\bar{A} = \int_{\mathbb{T}^d} \left(A(y)^{\mathrm{T}} + \nabla_y\chi(y)\, A(y)^{\mathrm{T}}\right)\mathrm{d}y.$$

实际上, 可将该表达式和 (12.3.1) 合并 (如平均) 可得到其他同样有效的 $\bar{A}$ 的表达式 (如对称). □

12.4 推　导

由于在方程 (12.2.1) 中出现一个小参数 ε, 通常会找一个在 ε 上进行幂级数展开的解

$$u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \cdots.$$

多尺度方法背后的基本思想是假设表达式里的所有项显式地依赖于 x 和 $y = x/\varepsilon$, 而且由于偏微分方程的系数是 x/ε 的周期函数, 则展开式中的所有项也应该是 x/ε 的周期函数, 从而假设解 u^ε 的形式如下:

$$u^\varepsilon(x) = u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(x, \frac{x}{\varepsilon}\right) + \cdots, \tag{12.4.1}$$

其中 $u_j(x, y)\,(j = 0, 1, \cdots)$ 在 y 上具有周期性.

变量 x 和 $y = x/\varepsilon$ 分别代表问题中的慢 (宏观) 尺度和快 (微观) 尺度. 当 $\varepsilon \ll 1$ 时, 变量 y 比 x 变化得更快, 并且在微观尺度下考虑问题时, 可将 x 看成是常数. 这就应用了尺度分离, 即将 x 和 y 看成是独立变量. 均匀化理论中主要问题之一就是证明当 $\varepsilon \to 0$ 时该假设的合理性, 见第 19 章.

实际上, $y = x/\varepsilon$ 暗含了关于 x 的偏导数变为

$$\nabla \to \nabla_x + \frac{1}{\varepsilon}\nabla_y.$$

换言之, 函数 $g^\varepsilon(x) := g\left(x, \frac{x}{\varepsilon}\right)$ 的全导数可表示如下:

$$\nabla g^\varepsilon(x) = \nabla_x g(x, y)\big|_{y=\frac{x}{\varepsilon}} + \frac{1}{\varepsilon}\, \nabla_y g(x, y)\big|_{y=\frac{x}{\varepsilon}},$$

其中记号 $h(x,y)|_{y=z}$ 表示当 $y=z$ 时, 函数 $h(x,y)$ 的值.

重新将微分算子

$$\mathcal{A}^{\varepsilon} := -\nabla \cdot (A(y)\nabla)$$

记为

$$\mathcal{A}^{\varepsilon} = \frac{1}{\varepsilon^2}\mathcal{A}_0 + \frac{1}{\varepsilon}\mathcal{A}_1 + \mathcal{A}_2, \tag{12.4.2}$$

其中

$$\mathcal{A}_0 := -\nabla_y \cdot (A(y)\nabla_y), \tag{12.4.3a}$$

$$\mathcal{A}_1 := -\nabla_y \cdot (A(y)\nabla_x) - \nabla_x \cdot (A(y)\nabla_y), \tag{12.4.3b}$$

$$\mathcal{A}_2 := -\nabla_x \cdot (A(y)\nabla_x). \tag{12.4.3c}$$

注意到这里定义的所有算子的系数都是 y 的周期函数. 令 $\mathcal{A}_0$ 在 $\mathbb{T}^d$ 上具有周期边界条件.

根据 (12.4.2), 方程 (12.2.1) 变为

$$\left(\frac{1}{\varepsilon^2}\mathcal{A}_0 + \frac{1}{\varepsilon}\mathcal{A}_1 + \mathcal{A}_2\right)u^{\varepsilon} = f, \quad (x,y) \in \Omega \times \mathbb{T}^d, \tag{12.4.4a}$$

$$u^{\varepsilon} = 0, \quad (x,y) \in \partial\Omega \times \mathbb{T}^d. \tag{12.4.4b}$$

将 (12.4.1) 代入 (12.4.4) 推得

$$\frac{1}{\varepsilon^2}\mathcal{A}_0 u_0 + \frac{1}{\varepsilon}(\mathcal{A}_0 u_1 + \mathcal{A}_1 u_0) + (\mathcal{A}_0 u_2 + \mathcal{A}_1 u_1 + \mathcal{A}_2 u_0) + O(\varepsilon) = f. \tag{12.4.5}$$

令方程 (12.4.5) 中 ε 等幂的系数等于零, 并且不计阶数高于 1 的项, 得到如下方程:

$$O\left(\frac{1}{\varepsilon^2}\right) : \mathcal{A}_0 u_0 = 0; \tag{12.4.6a}$$

$$O\left(\frac{1}{\varepsilon}\right) : \mathcal{A}_0 u_1 = -\mathcal{A}_1 u_0; \tag{12.4.6b}$$

$$O(1) : \mathcal{A}_0 u_2 = -\mathcal{A}_1 u_1 - \mathcal{A}_2 u_0 + f, \tag{12.4.6c}$$

其中 $u_j(x,y)$ 在第二参数上的周期为 1.

注意到在 y 上 $\mathcal{A}_0$ 是一个微分算子, 并且方程 (12.4.6b) 和 (12.4.6c) 中的 x 仅仅是一个参数. 由 (12.4.6a) 推得 $u_0(x,y) = u(x)$. 因此, 多尺度展开式中的首项独立于 y. 在剩下的两个形如 (12.2.3) 的方程中有 $v = v(x,y)$ 和 $h = h(x,y)$. 因此, x 是一个参数.

继续讨论方程 (12.4.6b), 其变为

$$\mathcal{A}_0 u_1 = \left(\nabla_y \cdot A^{\mathrm{T}}\right) \cdot \nabla_x u, \quad u_1(x, \cdot)\text{ 周期为 } 1, \quad \int_{\mathbb{T}^d} u_1 \mathrm{d}y = 0. \tag{12.4.7}$$

因为

$$\int_{\mathbb{T}^d} \left(\nabla_y \cdot A^{\mathrm{T}}\right) \cdot \nabla_x u \mathrm{d}y = \nabla_x u \cdot \int_{\mathbb{T}^d} \nabla_y \cdot A^{\mathrm{T}} \mathrm{d}y = 0,$$

根据收敛定理以及 $A(\cdot)$ 的周期性, 可知 (12.2.4) 满足可解性条件, 见注 7.13. 使用变量分离得到 (12.4.7) 的解为

$$u_1(x, y) = \chi(y) \cdot \nabla_x u(x). \tag{12.4.8}$$

将 (12.4.8) 代入到 (12.4.7) 中得到向量场 $\chi(y)$ 的单胞问题 (12.3.2), 称 $\chi(y)$ 为**一阶校正**. 注意: 系数的周期性暗含方程 (12.3.2) 右边在单胞上均值为零, 从而使单胞问题满足适定性. 通过令校正场均值为零 (条件 (12.3.4)) 来保证 (12.3.2) 解的唯一性.

现在考虑方程 (12.4.6c). 由 (12.2.4) 可知, 为了使此方程具有适定性, 方程右边在 $\mathbb{T}^d$ 上的均值必须为零. 由于已经假设函数 $f(x)$ 独立于 y, 则可解性条件暗含

$$\int_{\mathbb{T}^d} (\mathcal{A}_2 u_0 + \mathcal{A}_1 u_1) \mathrm{d}y = f. \tag{12.4.9}$$

方程 (12.4.9) 左边的第一项为

$$\begin{aligned}
\int_{\mathbb{T}^d} \mathcal{A}_2 u_0 \mathrm{d}y &= \int_{\mathbb{T}^d} -\nabla_x \cdot (A(y) \nabla_x u) \mathrm{d}y \\
&= -\nabla_x \cdot \left[\left(\int_{\mathbb{T}^d} A(y) \mathrm{d}y\right) \nabla_x u(x)\right] \\
&= -\left(\int_{\mathbb{T}^d} A(y) \mathrm{d}y\right) : \nabla_x \nabla_x u(x),
\end{aligned} \tag{12.4.10}$$

而且

$$\int_{\mathbb{T}^d} \mathcal{A}_1 u_1 \mathrm{d}y = \int_{\mathbb{T}^d} \left(-\nabla_y \cdot (A(y) \nabla_x u_1) - \nabla_x \cdot (A(y) \nabla_y u_1)\right) \mathrm{d}y =: I_1 + I_2$$

由周期性和注 7.13 得 $I_1 = 0$. 现在考虑 I_2.

$$\begin{aligned}
I_2 &= \int_{\mathbb{T}^d} -\nabla_x \cdot (A(y) \nabla_y u_1) \mathrm{d}y \\
&= -\int_{\mathbb{T}^d} A(y) : \nabla_x \nabla_y (\chi \cdot \nabla_x u) \mathrm{d}y \\
&= -\left(\int_{\mathbb{T}^d} \left(A(y) \nabla_y \chi(y)^{\mathrm{T}}\right) \mathrm{d}y\right) : \nabla_x \nabla_x u.
\end{aligned} \tag{12.4.11}$$

将 (12.4.11) 和 (12.4.10) 代到 (12.4.9) 中得到结论 12.1 的均匀化方程, 其中均匀化系数 $\bar{A}$ 由式 (12.3.1) 给出, 这就是推导过程.

12.5 简化方程的性质

本节中研究一些有效系数的基本特性. 特别地, 均匀化系数矩阵 $\bar{A}$ 是正定的, 这暗含均匀化系数算子是一致椭圆的, 从而也是适定的, 而且在均匀化条件下保持对称性, 即如果 $A(y)$ 具有对称性, 则均匀化矩阵也具有对称性. 还证明均匀化过程能产生各向异性的现象: 即使矩阵 $A(y)$ 是对角矩阵, 均匀化系数矩阵 $\bar{A}$ 也未必是对角矩阵.

为了研究均匀化系数矩阵, 找到一个 $\bar{A}$ 的替换式是非常有用的. 为此, 引入一个对所有函数 $\phi,\psi\in C^1(\mathbb{T}^d)$ 定义的双线性形式

$$a_1(\psi,\phi)=\int_{\mathbb{T}^d}\langle\nabla_y\phi, A(y)\nabla_y\psi\rangle\mathrm{d}y. \tag{12.5.1}$$

注意到这是与算子 $\mathcal{A}_0$ 有关的双线性形式.

$$\int_{\mathbb{T}^d}\phi\mathcal{A}_0\psi\mathrm{d}y=a_1(\phi,\psi),\quad \forall\phi,\psi\in C^1_{\mathrm{per}}(\mathbb{T}^d). \tag{12.5.2}$$

注意: 任何时候 A 是对称的, 双线性形式 $a_1(\cdot,\cdot)$ 也是对称的. 先找一个单胞问题的等价替换形式. 该替换形式在收敛形式上与第 7 章中介绍的椭圆型偏微分方程的弱形式密切相关. 在本节剩余部分中, 假设单胞问题的解是足够光滑的来证明随后计算的合理性. 假设校正场 $\chi(y)$ 的每个元素都是连续可微的, 并且具有周期性, $\chi_\ell(y)\in C^1_{\mathrm{per}}(\mathbb{T}^d)\ (\ell=1,\cdots,d)$.

回想 e_ℓ 定义为第 i 行为 $\delta_{i\ell}$ 的单位向量, 并且令 y_ℓ 为场 y 的第 ℓ 个元素. 注意到 $e_\ell=\nabla_y y_\ell$, 并且回想 $a_\ell=Ae_\ell$, 矩阵 A 的第 ℓ 列. 使用这两个基本事实可以得到下面有用的引理.

引理 12.3　单胞问题 (12.3.2) 可记为如下形式:

$$a_1(\phi,\chi_\ell+y_\ell)=0,\quad \forall\phi\in C^1_{\mathrm{per}}(\mathbb{T}^d),\ell=1,\cdots,d. \tag{12.5.3}$$

证明　由 (12.3.5) 推得

$$\mathcal{A}_0\chi_\ell=\nabla_y\cdot(Ae_\ell)=\nabla_y\cdot(A\nabla_y y_\ell)=-\mathcal{A}_0 y_\ell,$$

从而单胞问题可记为如下形式:

$$\mathcal{A}_0(\chi_\ell+y_\ell)=0,\quad l=1,\cdots,d,$$

其具有周期边界条件. 将前面的公式乘以一个任意函数 $\phi\in C^1_{\mathrm{per}}(\mathbb{T}^d)$. 应用注 7.13, 方程 (12.5.1) 以及方程 (12.5.2), 在单胞上积分得 (12.5.3). □

引理 12.3 给出了均匀化系数的替代公式, 表明在任何时候只要 $A(y)$ 具有对称性, 则 $\bar{A}$ 也具有对称性.

引理 12.4 有效矩阵 $\bar{A}$ 的元素如下:

$$\bar{a}_{ij} = a_1(\chi_j + y_j, \chi_i + y_i), \quad i, j = 1, \cdots, d. \tag{12.5.4}$$

特别地, $A(y)$ 的对称性暗含了 $\bar{A}$ 也具有对称性.

证明 首先注意到前面的定理暗含由于 $\chi_i(y) \in C^1_{\mathrm{per}}(\mathbb{T}^d)$, 使得

$$a_1(\chi_i, \chi_j + y_j) = 0, \quad \forall i, j = 1, \cdots, d. \tag{12.5.5}$$

现在由式 (12.3.1) 及 (12.5.5) 得

$$\begin{aligned}
\bar{a}_{ij} &= e_i \cdot \bar{A} e_j \\
&= \int_{\mathcal{Y}} \left(e_i \cdot A e_j + e_i \cdot A \nabla_y \chi^{\mathrm{T}} e_j\right) \mathrm{d}y \\
&= \int_{\mathcal{Y}} \left(\nabla_y y_i \cdot A \nabla_y y_j + \nabla_y y_i \cdot A \nabla_y \chi_j\right) \mathrm{d}y \\
&= \int_{\mathcal{Y}} \langle \nabla_y y_i, A(\nabla_y(y_j + \chi_j)) \rangle \mathrm{d}y \\
&= a_1(y_i, \chi_j + y_j) \\
&= a_1(y_i + \chi_i, \chi_j + y_j).
\end{aligned}$$

这就证明了 (12.5.4). 假设 $A(y) = A(y)^{\mathrm{T}}$, 这暗含双线性形式 $a_1(\cdot, \cdot)$ 具有对称性, 则

$$\bar{a}_{ij} = a_1(y_i + \chi_i, \chi_j + y_j) = a_1(y_j + \chi_j, \chi_i + y_i) = \bar{a}_{ji},$$

这表明均匀化矩阵具有对称性. □

下面将证明均匀化矩阵 $\bar{A}$ 是一致正定的, 这暗含均匀化方程是一个适定性椭圆型偏微分方程.

定理 12.5 均匀化系数矩阵 $\bar{A}$ 是一致正定的.

证明 定义 $\xi \in \mathbb{R}^d$ 为一个任意的向量, 需要证明存在一个常数 $\bar{\alpha} > 0$, 使得

$$\langle \xi, \bar{A}\xi \rangle \geqslant \bar{\alpha} |\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

由式 (12.5.4) 推得

$$\langle \xi, \bar{A}\xi \rangle = a_1(w, w),$$

其中 $w = \xi \cdot (\chi + y)$. 现在由 $A(y)$ 一致正定的定义得

$$a_1(w, w) \geqslant \alpha \int_{\mathbb{T}^d} |\nabla_y w|^2 \mathrm{d}y \geqslant 0$$

则是非负的.

为了证明通常 $\bar{A}$ 是正定的, 作如下讨论: 对一些 ξ, 假设

$$\langle \xi, \bar{A}\xi \rangle = 0.$$

由于 $\alpha > 0, \nabla_y w = 0$, 并且 $w = c$(c 是一个常向量), 则

$$\xi \cdot y = c - \xi \cdot \chi.$$

此方程右边以 1 为周期, 并在 y 上是连续的, 则左边同样有此性质. 只有当 $\xi = 0$ 时才出现这种情况, 这就证明了定理 12.5. □

定理 12.5 表明在均匀化过程中一致椭圆性得到了保留. 特别地, 这暗含由于均匀化方程是参数为常数的一致椭圆性偏微分方程, 因此, 其具有很好的适定性.

注 12.6　注意到均匀化不保持各向同性. 特别地, 即使扩散矩阵 A 仅有非零的对角元素, 通常均匀化扩散矩阵 $\bar{A}$ 也会有非零的非对角元素. 为此, 假设 $a_{ij} = 0(i \neq j)$, 则均匀化扩散矩阵的非对角元素由下面的公式给出 (这里没有求和缩写法):

$$\bar{a}_{ij} = \int_{\mathbb{T}^d} a_{ii} \frac{\partial \chi_j}{\partial y_i} \mathrm{d}y, \quad i \neq j.$$

这个表达式未必等于零, 并且会得到一个令人吃惊的结果, 即当微观结构变得越来越小时, 各向同性的复合材料表现出均匀化材料各向异性的特性. □

12.6　应　　用

本节列举两个例子来说明前面所述理论的应用, 在例子中可能会得到显式解. 从本质上来说, 一维情形仅是作了一般的设定, 即单胞问题可得到解析解, 并且可以得到有效扩散系数的显式. 在更高维的情形下, 只有当所研究问题的具体结构使得能够将均匀化系数的计算降低为一维问题时, 才能得到有效扩散系数的显式. 这种降维在层状材料下是可能的, 这就是将要讨论的第二个例子.

12.6.1　一维情形

设 $d = 1$ 且 $\Omega = [0, L]$, 则 Dirichlet 问题 (12.2.1a) 归结为一个两点边界值问题

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(a\left(\frac{x}{\varepsilon}\right)\frac{\mathrm{d}u^\varepsilon}{\mathrm{d}x}\right) = f, \quad x \in (0, L), \tag{12.6.1a}$$

$$u^\varepsilon(0) = u^\varepsilon(L) = 0. \tag{12.6.1b}$$

假设 $a(y)$ 是光滑的, 并且周期为 1, 还假设存在常数 $0 < \alpha \leqslant \beta < \infty$, 使得

$$\alpha \leqslant a(y) \leqslant \beta, \quad \forall y \in [0, 1], \tag{12.6.2}$$

并且假设 f 是光滑的.

单胞问题变为一个具有周期边界条件的常微分方程的边界值问题

$$-\frac{\mathrm{d}}{\mathrm{d}y}\left(a(y)\frac{\mathrm{d}\chi}{\mathrm{d}y}\right)=\frac{\mathrm{d}a(y)}{\mathrm{d}y},\quad y\in(0,1), \tag{12.6.3a}$$

$$\chi\text{的周期为 }1,\quad \int_0^1\chi(y)\mathrm{d}y=0. \tag{12.6.3b}$$

由于 $d=1$, 由 (12.3.1) 的一维情形仅得到一个有效系数, 即

$$\bar{a}=\int_0^1\left(a(y)+a(y)\frac{\mathrm{d}\chi(y)}{\mathrm{d}y}\right)\mathrm{d}y=\left\langle a(y)\left(1+\frac{\mathrm{d}\chi(y)}{\mathrm{d}y}\right)\right\rangle. \tag{12.6.4}$$

在本章剩余部分, 采用记号

$$\langle f(y)\rangle:=\int_{\mathbb{T}^d}f(y)\,\mathrm{d}y$$

来表示在 $\mathbb{T}^d$ 上的平均.

方程 (12.6.3a) 有精确解. 在 0 到 y 上积分得

$$a(y)\frac{\mathrm{d}\chi}{\mathrm{d}y}=-a(y)+c_1, \tag{12.6.5}$$

其中常数 c_1 是不确定的. 由于不等式 (12.6.2) 暗含 a 是严格正定的, 就可以用 $a(y)$ 除以 (12.6.5), 然后再一次在 0 到 y 上积分得

$$\chi(y)=-y+c_1\int_0^y\frac{1}{a(y)}\mathrm{d}y+c_2.$$

由 $\chi(y)$ 是一个周期函数来确定常数 c_1, 则 $\chi(0)=\chi(1)$, 并且推得

$$c_1=\frac{1}{\displaystyle\int_0^1\frac{1}{a(y)}\mathrm{d}y}=\left\langle a(y)^{-1}\right\rangle^{-1}.$$

由 (12.6.5) 有

$$1+\frac{\mathrm{d}\chi}{\mathrm{d}y}=\frac{1}{\left\langle a(y)^{-1}\right\rangle a(y)}.$$

(注意到在计算 $\bar{a}$ 时, 对 c_2 不作要求.) 将这个表达式代入方程 (12.6.4) 得

$$\bar{a}=\left\langle a(y)^{-1}\right\rangle^{-1}. \tag{12.6.6}$$

这就是一维情形下均匀化系数的计算公式. 显然, 即使在这种简单的一维情形下, 仅仅在微观结构的一段时间上求不均匀系数的平均是不能得到均匀化系数的. 更确切地说, 均匀化系数是不均匀化系数倒数 (调和平均数) 平均数的倒数. 显然, 由调和平均数 (12.6.6) 给出的均匀化系数受 $a(y)$ 平均值的约束, 见练习第 12 题.

12.6.2 层状材料

考虑二维情形下的问题 (12.2.1), 其满足假设 (12.2.2). 假设域 $\Omega \subset \mathbb{R}^2$ 代表层状材料, 材料性质仅在一个方向上变化. 因此, 系数 $A(y)$ 是关于一个变量的函数. 对 $y=(y_1,y_2)^{\mathrm{T}}$ 有

$$a_{ij}=a_{ij}(y_1), \quad i,j=1,2. \tag{12.6.7}$$

事实上, 系数是 y_1 的函数暗含单胞问题 (12.3.2) 右边只是 y_1 的函数. 因此, 单胞问题的解也仅仅是 y_1 的函数, 形式如下:

$$\chi_\ell=\chi_\ell(y_1), \quad \ell=1,2. \tag{12.6.8}$$

将式 (12.6.8) 代入 (12.3.2), 单胞问题变为

$$-\frac{\mathrm{d}}{\mathrm{d}y_1}\left(a_{11}(y_1)\frac{\mathrm{d}\chi_\ell(y_1)}{\mathrm{d}y_1}\right)=\frac{\mathrm{d}a_{1\ell}(y_1)}{\mathrm{d}y_1}, \quad \ell=1,2, \tag{12.6.9}$$

其具有周期边界条件. 类似地, 均匀化参数的式 (12.3.1) 变为

$$\bar{a}_{ij}=\int_0^1\left(a_{ij}(y_1)+a_{i1}(y_1)\frac{\mathrm{d}\chi_j(y_1)}{\mathrm{d}y_1}\right)\mathrm{d}y_1, \quad i,j=1,2. \tag{12.6.10}$$

现在求解方程 (12.6.9). 这些方程都是常微分方程, 可以用与前面小节中求解一维情形一样的方法来求解这些方程. 为此, 将其在 0 到 y 上积分, 并除以 $a_{11}(y_1)$ 得

$$\frac{\mathrm{d}\chi_\ell}{\mathrm{d}y_1}=-\frac{a_{1\ell}}{a_{11}}+c_1\frac{1}{a_{11}}, \quad \ell=1,2, \tag{12.6.11}$$

其中常数 c_1 待定. 必须分别考虑 $\ell=1$ 和 $\ell=2$ 的情形. 首先考虑 $\ell=1$, 此时方程简化为

$$\frac{\mathrm{d}\chi_1}{\mathrm{d}y_1}=-1+c_1\frac{1}{a_{11}},$$

这个方程与在 12.6.1 小节中考虑过的方程完全一致, 则有

$$\frac{\mathrm{d}\chi_1}{\mathrm{d}y_1}=-1+\frac{1}{\left\langle a_{11}(y)^{-1}\right\rangle a_{11}(y)}. \tag{12.6.12}$$

现在考虑 $\ell=2$ 时的方程 (12.6.11).

$$\frac{\mathrm{d}\chi_2}{\mathrm{d}y_1}=-\frac{a_{12}}{a_{11}}+c_1\frac{1}{a_{11}},$$

再次对其积分, 然后通过令 $\chi_2(y_1)$ 具有周期性来确定 c_1. 最终结果为

$$\frac{\mathrm{d}\chi_2(y_1)}{\mathrm{d}y_1}=-\frac{a_{12}(y_1)}{a_{11}(y_1)}+\frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle}\frac{1}{a_{11}(y_1)}. \tag{12.6.13}$$

现在就能够计算均匀化系数了. 首先计算 $\bar{a}_{11}$, 计算过程与一维情形一样.

$$\bar{a}_{11} = \left\langle a_{11}(y_1)^{-1} \right\rangle^{-1}. \tag{12.6.14}$$

继续计算 $\bar{a}_{12}$, 将 (12.6.13) 代到 (12.6.10) 中, 其中 $i=1, j=2$ 得

$$\begin{aligned}
\bar{a}_{12} &= \int_0^1 \left(a_{12}(y_1) + a_{11}(y_1) \frac{\mathrm{d}\chi_2(y_1)}{\mathrm{d}y_1} \right) \mathrm{d}y \\
&= \int_0^1 \left(a_{12}(y_1) + a_{11}(y_1) \left(-\frac{a_{12}(y_1)}{a_{11}(y_1)} + \frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle} \frac{1}{a_{11}(y_1)} \right) \right) \mathrm{d}y \\
&= \int_0^1 \left(a_{12}(y_1) - a_{12}(y_1) + \frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle} \right) \mathrm{d}y \\
&= \frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle},
\end{aligned}$$

则

$$\bar{a}_{12} = \left\langle \frac{a_{12}(y_1)}{a_{11}(y_1)} \right\rangle \left\langle a_{11}^{-1}(y_1) \right\rangle^{-1}. \tag{12.6.15}$$

类似地,

$$\bar{a}_{21} = \left\langle \frac{a_{21}(y_1)}{a_{11}(y_1)} \right\rangle \left\langle a_{11}^{-1}(y_1) \right\rangle^{-1}. \tag{12.6.16}$$

最后考虑 $\bar{a}_{22}$.

$$\begin{aligned}
\bar{a}_{22} &= \int_0^1 \left(a_{22}(y_1) + a_{21}(y_1) \frac{\mathrm{d}\chi_2(y_1)}{\cdot\, \mathrm{d}y_1} \right) \mathrm{d}y \\
&= \int_0^1 \left(a_{22}(y_1) + a_{21}(y_1) \left(-\frac{a_{12}(y_1)}{a_{11}(y_1)} + \frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle} \frac{1}{a_{11}(y_1)} \right) \right) \mathrm{d}y \\
&= \int_0^1 \left(a_{12}(y_1) - \frac{a_{12}(y_1)a_{21}(y_1)}{a_{11}(y_1)} + \frac{a_{21}(y_1)}{a_{11}(y_1)} \frac{\langle a_{12}(y_1)/a_{11}(y_1)\rangle}{\langle a_{11}^{-1}(y_1)\rangle} \right) \mathrm{d}y \\
&= \left\langle \frac{a_{21}(y_1)}{a_{11}(y_1)} \right\rangle \left\langle \frac{a_{12}(y_1)}{a_{11}(y_1)} \right\rangle \langle a_{11}^{-1}(y_1)\rangle^{-1} + \left\langle a_{22}(y_1) - \frac{a_{12}(y_1)a_{21}(y_1)}{a_{11}(y_1)} \right\rangle.
\end{aligned}$$

结果为

$$\bar{a}_{22} = \left\langle \frac{a_{21}(y_1)}{a_{11}(y_1)} \right\rangle \left\langle \frac{a_{12}(y_1)}{a_{11}(y_1)} \right\rangle \left\langle a_{11}^{-1}(y_1) \right\rangle^{-1} + \left\langle a_{22}(y_1) - \frac{a_{12}(y_1)a_{21}(y_1)}{a_{11}(y_1)} \right\rangle. \tag{12.6.17}$$

显然, 由式 (12.6.14)~(12.6.17) 可知, 均匀化系数在一个非常复杂、高度非线性的方式下依赖于初始值.

12.7 讨论和参考

在 19 世纪 70 年代, 许多研究者都对多尺度方法的发展作出了贡献, 其中比较出名的学者有 Keller, Babuska, Sanchez-Palenzia, Bensoussan, Lions, Papanicolaou 以及其他的研究者, 见文献 [23~26,91, 158, 159] 及其参考文献. 阐述多尺度方法的第一个系统可在文献 [33] 中找到, 其参考文献中列出了更早关于多尺度方法的文献, 也可见文献 [279]; 在多尺度方法发展以前, 就已对具有快速振荡系数的椭圆型偏微分方程的严格收敛结论进行了证明, 见文献 [73, 296] 以及教材文献 [153]. 然而, 多尺度方法可以广泛地应用到不同问题中去. 与之相反, 将严格的结论应用于不同情形时, 在不同的偏微分方程、Markov 链、常微分方程、随机微分方程中的证明是大相径庭的 (见本书的第三部分). 然而, 在大多数情况下, 一个恰当的 Poisson 方程 (单胞问题) 在分析中扮演着很重要的角色.

书 [296] 中没有使用多尺度方法来研究一维问题 (见 12.6.1 小节). 在一维情形下, 即使在非周期条件下使用多尺度方法来推导均匀化方程也是有可能的, 见文献 [143, 66, ch.5]. 就非周期性快速振荡系数来说, 均匀化方程是一个二阶一致椭圆型偏微分方程. 然而, 由 H 和 Γ 收敛理论很自然地得到这个结论, 而不需要用多尺度展开, 见文献 [296, 308]. 在非周期性、确定性的均匀化背景下, 均匀化系数不能通过相应的 Poisson 方程的解来表示. 此时, 得到均匀化系数的界限是最好的.

没有使用多尺度方法, Murat 和 Tartar 严格推导出了层状材料 (见 12.6.2 小节) 的均匀化方程, 见文献 [232] 及其原稿的参考文献. 在 12.6.2 小节中讨论的二维情形可以很容易地延伸到 d 维情形, 即 $d \geqslant 2$, 其中 $a_{ij}(y) = a_{ij}(y_1)\,(i,j=1,\cdots,d)$, 见文献 [232].

椭圆边值问题 (12.2.1) 是一个 Dirichlet 问题. 然而, 检查 12.4 节中所给出的分析可知, 边界条件在推导均匀化方程的过程中没有起到任何作用. 特别地, 用于推导均匀化方程的双尺度展开 (12.4.1) 中没有包含任何关于所研究问题边界条件的信息. 实际上, 在均匀化过程中, 边界条件变得有些无关紧要了. 正因如此, 进行同样的计算使得可以得到关于 Neumann 或者是混合边界条件的均匀化方程. 因为推导均匀化方程是基于形如 (12.2.3) 的局部问题的分析, 所以就会很自然地得到上述结果. 这个局部问题不能真正地看到边界 —— 这是尺度分离的主要特性.

然而, 当想要证明均匀化理论时, 边界条件就变得很重要了. 事实上, 双尺度展开 (12.4.1) 不满足偏微分方程的边界条件, 确切地讲, 只有 $O(\varepsilon)$, 从而只能引入边界层 [143, ch.3]①. 边界层影响 $u^{\varepsilon}(x)$ 收敛于 $u(x)$ 的收敛率 (当 $\varepsilon \to 0$ 时). 通过修

① 在所有的奇异扰动问题中, 一般都会出现边界初始层. 进一步的说明见本书第二部分章节中所列的参考文献以及文献 [146, 161].

正双尺度展开 (12.4.1), 增加附加项可以解决上述问题, 增加的附加项抵消了边界层的影响. 因为当离开边界时, 其影响呈指数衰减, 从而使得它们不能影响内部区域的解, 详见 [27].

文献 [33] 和文献 [66] 中进一步说明了注 12.2 中讨论的情形. 在证明有效扩散张量的不同性质时, 要用到有效扩散张量的不同表达式.

从连续力学的观点来讲, 使用均匀化方法可得到复合材料的宏观本构关系. 使用均匀化理论已经推导出了不同类型复合材料的宏观本构关系, 见文献 [46, 108]. [133, 230] 中给出一种替换方法. 在优秀专著 [229] 中给出了有关复合材料的理论.

在 12.4 节中分析的 Dirichlet 问题中, 假定了矩阵 $A^{\varepsilon}(x)$ 仅依赖于微观尺度, 即

$$A^{\varepsilon}(x) = A\left(\frac{x}{\varepsilon}\right),$$

其中 $A(y)$ 是一个以 1 为周期的矩阵值函数. 然而, 多尺度方法也可应用于系数明显依赖于宏观以及微观尺度的情形,

$$A^{\varepsilon}(x) = A\left(x, \frac{x}{\varepsilon}\right),$$

其中 $A(x,y)$ 在 y 上以 1 为周期, 并且在 x 上是光滑的. 当系数有此形式时, 就称之为局部周期或非均匀周期. 应用类似于 12.4 节中给出的分析, 可以得到 Dirichlet 问题的均匀化方程

$$-\nabla \cdot (A^{\varepsilon} \nabla u^{\varepsilon}) = f, \quad x \in \Omega, \tag{12.7.1a}$$

$$u^{\varepsilon} = 0, \quad x \in \partial\Omega, \tag{12.7.1b}$$

其中 $A^{\varepsilon}(x) = A(x, x/\varepsilon)$. 现在均匀化系数 $\bar{A}$ 是 x 的函数,

$$-\nabla \cdot \left(\bar{A} \nabla u\right) = f, \quad x \in \Omega, \tag{12.7.2a}$$

$$u = 0, \quad x \in \partial\Omega, \tag{12.7.2b}$$

并且由于 $A = A(x,y)$, 使得单胞问题是 x 的参数,

$$-\nabla_y \cdot \left(\nabla_y \chi A^{\mathrm{T}}\right) = \nabla_y \cdot A^{\mathrm{T}}, \quad y \in \mathbb{T}^d. \tag{12.7.3}$$

由如下公式给出均匀化系数:

$$\bar{A}(x) = \int_{\mathbb{T}^d} \left(A(x,y) + A(x,y) \nabla_x \chi(x,y)^{\mathrm{T}}\right) \mathrm{d}y. \tag{12.7.4}$$

强调如下事实: 前面两个方程中出现的 "宏观变量" x 是一个参数. 结果, 为了计算有效系数, 需要求解单胞问题 (12.7.3), 并且在 (12.7.4) 中对所有 $x \in \Omega$ 求解积分.

多尺度方法也可用于解决半线性椭圆型偏微分方程, 该方程具有如下形式的快速振荡系数方程:

$$-\nabla\cdot\left(A^{\varepsilon}\nabla u^{\varepsilon}\right)=f\left(u^{\varepsilon}\right),\quad x\in\Omega, \tag{12.7.5a}$$

$$u^{\varepsilon}=0,\quad x\in\partial\Omega, \tag{12.7.5b}$$

均匀化方程的形式为

$$-\nabla\cdot\left(\bar{A}\nabla u\right)=f\left(u\right),\quad x\in\Omega, \tag{12.7.6a}$$

$$u=0,\quad x\in\partial\Omega, \tag{12.7.6b}$$

其中 $\bar{A}$ 定义如 (12.3.1).

在 12.2 节中得到了 Dirichlet 问题 (12.2.1) 双尺度展开的前两项, 在展开式中, 第二项与一个 x 的函数和第一项的梯度的乘积成比例, 该展开式为均匀化方程的解, 即

$$u_1\left(x,\frac{x}{\varepsilon}\right)=\chi\left(\frac{x}{\varepsilon}\right)\cdot\nabla_x u\left(x\right)+\hat{u}_1\left(x\right), \tag{12.7.7}$$

其中 $\chi\left(y\right)$ 为单胞问题的解. 也可求解更高阶的方程, 并得到双尺度展开式中更高阶的项. 例如, 可以解方程 (12.4.6), 并且计算展开式中的第三项 $u_2\left(x,y\right)$,

$$u_2\left(x,y\right)=\Theta\left(y\right):\nabla_x\nabla_x u\left(x\right)+\hat{u}_2\left(x\right), \tag{12.7.8}$$

其中二阶修正域 $\Theta\left(y\right)$ 是一个满足下面边界值问题的矩阵值函数:

$$\mathcal{A}_0\Theta=B, \tag{12.7.9}$$

$B\left(y\right)$ 由下式给出:

$$B\left(y\right):=-\bar{A}+A\left(y\right)+A\left(y\right)\nabla_y\chi\left(y\right)^{\mathrm{T}}+\nabla_y\chi\left(y\right)A\left(y\right)+\chi\left(y\right)\otimes\left(\nabla_y\cdot A\left(y\right)^{\mathrm{T}}\right).$$

所有更高阶方程形式如下:

$$\mathcal{A}_0u_{k+2}=-\mathcal{A}_1u_{k+1}-\mathcal{A}_0u_k,\quad k=1,2,\cdots.$$

这就证明了 $u_k\left(x\right)$ 与 $u\left(x\right)$ 第 k 阶导数成正比, 见 [27].

多尺度方法也可推广到存在 k 阶尺度的情形下, 即矩阵 $A^{\varepsilon}\left(x\right)$ 形式为

$$A^{\varepsilon}\left(x\right)=A\left(\frac{x}{\varepsilon},\frac{x}{\varepsilon^2},\cdots,\frac{x}{\varepsilon^k}\right),$$

并且 A 在其所有参数上是 1 周期的. 这就是重复均匀化, 见文献 [33,sec.1.8]. 有关一般情况下重复均匀化的严格分析见文献 [8]. 最近已经发现重复均匀化可应用于解决流体中的平流以及被动示踪的扩散问题, 详见文献 [219, 220, 253]. 当问题中存

在无穷多的尺度, 而没有尺度分离时, 均匀化结论就不成立, 此时均匀化系数可以是 0, 见文献 [16].

通常不可能求得均匀化系数的解析解. 的确, 这要求解出单胞问题以及计算 (12.3.1) 中的积分. 在大多数情况下, 只能得到这些问题的数值解. 然而, 得到有效系数值的范围是有可能的. 已经有不同方法来求得有效系数值范围, 如可以得到均匀化系数的变化特征, 有关这方面的不同结论见文献 [107, 229, 311]. 这些方法中的许多可应用于非周期背景下.

在 A^ε 和 f^ε 中, 对 ε 依赖性进行不同假设, 很容易将本章得到的方法延伸到初值 (或边界值) 问题上. 例如, 下面的抛物线型偏微分方程:

$$\frac{\partial u^\varepsilon}{\partial t} - \nabla \cdot (A^\varepsilon \nabla u^\varepsilon) = f^\varepsilon, \quad \Omega \times (0, T), \tag{12.7.10a}$$

$$u^\varepsilon = 0, \quad \partial\Omega \times (0, T), \tag{12.7.10b}$$

$$u^\varepsilon = u_{in}(x), \quad \bar{\Omega} \times \{0\}. \tag{12.7.10c}$$

当进化偏微分方程的系数在时间和空间上是振荡的, 即 $A^\varepsilon = A\left(x/\varepsilon, t/\varepsilon^k\right), k > 0$(矩阵值函数 $A(y, \tau)$ 在 y 和 τ 上是 1 周期的) 时, 就会产生一个有趣的时间依赖解. 这就意味着必须引入两个快速变量 $y = x/\varepsilon$ 和 $\tau = t/\varepsilon^k$. 有关具有时空依赖性系数的均匀化进化方程的更多信息见文献 [33,ch.3]. 将在第 11, 13 和 14 章中使用多尺度方法研究抛物线型偏微分方程的均匀化.

也可研究形式如下的双曲线 (波) 方程的均匀化问题:

$$\frac{\partial^2 u^\varepsilon}{\partial t^2} - \nabla \cdot (A^\varepsilon \nabla u^\varepsilon) = f, \quad \Omega \times (0, T), \tag{12.7.11a}$$

$$u^\varepsilon = 0, \quad \partial\Omega \times (0, T), \tag{12.7.11b}$$

$$u^\varepsilon = u_{\text{in}}, \quad \bar{\Omega} \times \{0\}, \tag{12.7.11c}$$

$$\frac{\partial u^\varepsilon}{\partial t} = v_{\text{in}}(x), \quad \bar{\Omega} \times \{0\}. \tag{12.7.11d}$$

可应用多尺度方法得到均匀化方程, 该方程是一个具有常系数, 并且初值和边界值条件相同的波动方程. 然而, 其与抛物线型偏微分方程有一个根本的区别: 对于抛物线型的问题, 不论初值如何选取, 消散使得解在首项生成子 $\mathcal{L}_0$ 的零空间附近. 对于波动方程, 这种情形不会发生, 并且有必要选取接近期望子空间的初值. 本书中不对波动方程的均匀化进行研究, 有兴趣的读者可参见文献 [33, ch.2, 47, 66,ch.12, 160]. 对于具有多尺度的 Schrödinger 方程的类似问题可见文献 [316]. [10] 中给出 Schrödinger 方程的均匀化理论及其与有效理论的联系.

在周期性条件下, 使用谱方法能直接得到均匀化方程的数值解. 另一方面, 当 ε 很小时, 得到初始边界值问题 (12.2.1) 的数值解是很困难的. 在过去的几年中, 已

经出现了一些特殊的方法, 这些方法都是以这样或那样的形式基于均匀化, 关于这个话题见文献 [2, 9, 52, 61, 76, 82, 89, 90, 92, 145, 231] 及其参考文献. 对具有多尺度结构的椭圆型偏微分方程进行有限单元法分析, 以及在地球物理学中的应用在文献 [60, 88, 145, 146] 中进行了讨论. 在文献 [3, 141, 215, 216, 286, 287] 中给出了服从随机力或具有随机系数的椭圆型偏微分方程的数值解法.

12.8 练　习

1. 当系数矩阵 $A(y)$ 在每个方向上具有不同周期时,

$$A(y+\lambda_k e_k)=A(y), \quad k=1,\cdots,$$

其中 $\lambda_k>0(k=1,\cdots,d)$. 考虑 (12.2.1) 的均匀化问题, 写出均匀化系数的公式.

2. 考虑问题 (12.2.1) 的双尺度展开 (12.4.1). 本章中用双尺度展开计算了前三项: u_0 是均匀化方程的解, (12.7.7) 和 (12.7.8) 分别给出了 u_1 和 u_2. 证明 u_2 的表达式以及高阶单胞问题 (12.7.9) 形式的正确性.

3. 在 d 维复合材料上考虑 Dirichlet 问题 (12.2.1), 即

$$a_{ij}(y)=a_{ij}(y_1), \quad 在\ y_1\ 上周期为\ 1, i, j=1,\cdots,d.$$

在 12.6.2 小节中, 当 $d=2$ 时求解此问题. 现在求解相关单胞问题, 并得到 $d\geqslant 3$ 或任意维时均匀化系数的计算公式.

4. 在一维情形下, 考虑二阶一致椭圆型偏微分方程的均匀化问题, 即在 12.6.1 小节中研究的问题.

(a) 当

$$a(y)=\begin{cases} a_1, & y\in\left[0,\dfrac{1}{2}\right], \\ a_2, & y\in\left(\dfrac{1}{2},1\right] \end{cases}$$

时, 计算 $\bar{a}$, 其中 a_1 和 a_2 是正常数;

(b) 当

$$a(y)=\frac{1}{2+\sin(2\pi y)}$$

时, 计算 $\bar{a}$.

5. 在 d 维各向同性材料上考虑 Dirichlet 问题 (12.2.1), 即

$$a_{ij}(y)=a(y)\delta_{ij}, \quad 周期为\ 1, i, j=1,\cdots,d,$$

其中 δ_{ij} 代表 Kronecker's delta 符号.

(a) 使用 $A(y)$ 的特殊结构来尽可能地简化单胞问题;

(b) 设 $d=2$, 并假设 $a(y)$ 形式如下:

$$a(y)=Y_1(y_1)\,Y_2(y_2)\,.$$

求解单胞问题的两个部分, 并且得到均匀化系数的计算公式. (提示: 使用变量分离.)

6. 考虑边界值问题 (12.7.1). 假设 $A^{\varepsilon}=A(x,x/\varepsilon)$, 其中 $A(x,y)$ 是光滑的、在 y 上周期为 1, 并且一致椭圆, f 是光滑的. 使用多尺度方法得到均匀化方程 (12.7.2) 的一般形式以及均匀化系数 (12.7.4) 的计算公式. 证明 12.5 节中的结论仍然成立.

7. 考虑 Dirichlet 问题

$$-\nabla\cdot\left(A\left(\frac{x}{\varepsilon},\frac{x}{\varepsilon^2}\right)\nabla u^{\varepsilon}\right)=f,\quad x\in\Omega, \tag{12.8.1a}$$

$$u^{\varepsilon}(x)=0,\quad x\in\partial\Omega, \tag{12.8.1b}$$

其中系数 $A(y,z)$ 在 y 和 z 上的周期为 1. 使用三尺度展开

$$u^{\varepsilon}(x)=u_0\left(x,\frac{x}{\varepsilon},\frac{x}{\varepsilon^2}\right)+\varepsilon u_1\left(x,\frac{x}{\varepsilon},\frac{x}{\varepsilon^2}\right)+\varepsilon^2u_2\left(x,\frac{x}{\varepsilon},\frac{x}{\varepsilon^2}\right)+\cdots$$

来推导一个有效均匀化方程, 以及均匀化系数的计算公式和单胞问题.

8. 先令 $z=y/\varepsilon$, 运用均匀化方法重复前面的练习, 然后对 y 做同样的练习.

(a) 将 y 看成是一个参数, 对下面的方程作均匀化处理:

$$-\nabla\cdot\left(A\left(y,\frac{y}{\varepsilon}\right)\nabla u^{\varepsilon}\right)=f,\quad x\in\Omega, \tag{12.8.2a}$$

$$u^{\varepsilon}(x)=0,\quad x\in\partial\Omega; \tag{12.8.2b}$$

(b) 均匀化下面的方程:

$$-\nabla\cdot\left(\bar{A}\left(\frac{x}{\varepsilon}\right)\nabla\bar{u}^{\varepsilon}\right)=f,\quad x\in\Omega, \tag{12.8.3a}$$

$$u^{\varepsilon}(x)=0,\quad x\in\partial\Omega, \tag{12.8.3b}$$

其中 $\bar{A}(y)$ 由前面一节中推出的表达式得到.

9. 通过将多尺度方法应用到热方程 (12.7.10) 上, 推导均匀化方程、单胞问题以及均匀化系数的计算公式, 其中 $A^{\varepsilon}=A(x/\varepsilon)$.

10. 考虑初始边界值问题 (12.7.10), 其中 $A^{\varepsilon}=A\left(x/\varepsilon,t/\varepsilon^k\right)$. 解释为什么通常示波器的周期为 $k=2$, 实现当 $k=1,2,3$ 时的均匀化过程①.

11. 由 (12.7.11) 使用多尺度方法推导均匀化方程.

12. 在 (12.6.2) 的条件下, 证明方程 (12.6.1) 的均匀化系数 $\bar{a}$ 与 $a(y)$ 有同样的上下界,

$$\alpha\leqslant\bar{a}\leqslant\beta,$$

而且通过 $a(y)$ 的平均

$$\bar{a}\leqslant\langle a(y)\rangle\,,$$

① 使用多尺度方法推导均匀化方程详见文献 [33,ch.3, 253].

证明其是有界的.

13. 证明通过对方程 (12.7.5) 进行均匀化可得有效方程 (12.7.6).

14. 设 $A : \mathbb{T}^d \to \mathbb{R}^{d\times d}$ 是光滑的, 并且具有周期性, 考虑特征值问题

$$-\nabla\cdot(A^\varepsilon\nabla u^\varepsilon)=\lambda^\varepsilon u^\varepsilon,\quad x\in\Omega,$$

$$u^\varepsilon=0,\quad x\in\partial\Omega,$$

其中 $A^\varepsilon = A(x/\varepsilon)$. 当 $\varepsilon\to 0$ 时, 使用多尺度展开来找到特征值问题的一个近似.

15. (a) 考虑特征值问题

$$-\Delta u^\varepsilon+\frac{1}{\varepsilon}V^\varepsilon u^\varepsilon=\lambda^\varepsilon u^\varepsilon,\quad x\in\Omega,$$

$$u^\varepsilon=0,\quad x\in\partial\Omega.$$

假设 $V : \mathbb{T}^d \to \mathbb{R}$ 是光滑的, 并且具有周期性, 以及

$$\int_{\mathbb{T}^d} V(y)\mathrm{d}y=0,$$

并且 $V^\varepsilon(x)=V(x/\varepsilon)$. 当 $\varepsilon\to 0$ 时, 使用多尺度展开来找到特征值问题的一个近似.

(b) 所得到的特征值比当 $V\equiv 0$ 时得到特征值大还是小?

第 13 章　抛物型偏微分方程的均匀化

13.1　引　　言

本章使用多尺度方法研究抛物型偏微分方程解的长时间行为. 使用的方法几乎与第 11 章中研究随机微分方程的均匀化所用的方法一致. 在本章末尾, 这种联系会变得更加明显.

在 13.2 节中给出了将要研究的完全方程. 13.3 节中包含了运用 13.4 节中的多尺度方法推导出的简化方程. 13.5 节列出了简化方程的不同特性. 在 13.6 节中研究了一般理论在两方面的应用, 一个是梯度流 (见 13.6.1 小节), 一个是无散流 (见 13.6.2 小节). 在 13.7 节中给出了抛物型偏微分方程均匀化和随机微分方程不对称问题二者之间的联系. 对本章内容的延伸及其参考文献见 13.8 节.

13.2　完 全 方 程

研究下面初值 (Cauchy) 问题:

$$\frac{\partial u}{\partial t} = b \cdot \nabla u + D\Delta u, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}^+, \tag{13.2.1a}$$

$$u = u_{\text{in}}, \quad (x,t) \in \mathbb{R}^d \times \{0\}, \tag{13.2.1b}$$

其中 $D > 0$. 在分析中, 假设向量 $b(x)$ 是光滑的, 在空间所有方向上的周期为 1, 而且假设初始条件是慢慢改变的, 因此,

$$u_{\text{in}}(x) = g^{\varepsilon}(x) := g(\varepsilon x), \tag{13.2.2}$$

其中 $0 < \varepsilon \leqslant 1$. 由于初值是慢慢变化的, 则解也会慢慢改变, 所以通常通过大空间和时间尺度来考察偏微分方程 (13.2.1) 的有效行为. 就像本章将要给出的那样, 如果在恰当意义上向量场 b 的均值为零, 则 u 的有效行为就是一个纯扩散行为.

为此, 通过重新标定尺度来定义变量 x, t,

$$x = \varepsilon^{-1}x, \quad t \to \varepsilon^{-2}t, \tag{13.2.3}$$

并重新标记 u 为 u^{ε} 来突出重新标定了的尺度. 无论何时由 b 引起的对流影响达到平衡, 这种特殊的时空尺度标定 (通常称为**扩散尺度**) 都是恰当的, 则在一个比空间

更长的尺度上标定时间, 并寻找纯扩散影响是合适的. 本节末会对 b 达到平衡这种情况进行明确的论述.

重新标定场 $u^\varepsilon(x,t)$ 满足方程

$$\frac{\partial u^\varepsilon}{\partial t}=\frac{1}{\varepsilon}b^\varepsilon\cdot\nabla u^\varepsilon+D\Delta u^\varepsilon,\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+, \tag{13.2.4a}$$

$$u^\varepsilon=g,\quad (x,t)\in\mathbb{R}^d\times\{0\}, \tag{13.2.4b}$$

其中 $b^\varepsilon(x)=b(x/\varepsilon)$. 该方程将是本章要研究的对象.

定义算子

$$\mathcal{L}_0=b(y)\cdot\nabla_y+D\Delta_y, \tag{13.2.5}$$

其在 $[0,1]^d$ 上具有周期边界条件, 并且其 L^2 伴随 L_0^* 也具有周期边界条件. 称 D 为**分子扩散系数**. 注意到 $\mathcal{L}_0$ 是 Markov 过程 $y(t)$ 的生成子, 其中 $y(t)$ 是如下随机微分方程在单位胞体 $\mathbb{T}^d$ 上的解:

$$\frac{\mathrm{d}y}{\mathrm{d}t}=b(y)+\sqrt{2D}\frac{\mathrm{d}W}{\mathrm{d}t}.$$

因此, 通常定义**不变分布**$\rho(y)$ 为下面伴随方程的稳态解:

$$\mathcal{L}_0^*\rho=0. \tag{13.2.6}$$

由定理 6.16 知该方程存在唯一解, 并能够规范化, 通过选择规范化形式使解是正定的, 则根据

$$\int_{\mathbb{T}^d}\rho(y)\mathrm{d}y=1$$

将解规范化到 (13.2.6). 注意到这样在 $\mathbb{T}^d$ 上就使得测度 $\mu(\mathrm{d}y)=\rho(y)\,\mathrm{d}y$ 变为一个概率测度.

为了推导 (13.2.4) 的均匀化方程需要研究如下形式的方程:

$$-\mathcal{L}_0v=h \tag{13.2.7}$$

其具有周期边界条件, 并且 h 是一个 y 的光滑周期函数. 显然, 满足定理 7.9 的假设, 因此, 算子 $\mathcal{L}_0$ 满足 Fredholm 性质. 特别地, 这就暗示 $\mathcal{L}_0$ 具有一个一维零空间, 在 y 上包含常数. 正如前面所述, 同样 $\mathcal{L}_0^*$ 有一个一维零空间且由 ρ 张成, 而且方程 (13.2.7) 有解的充要条件是方程的右边关于不变分布满足如下条件:

$$\int_{\mathbb{T}^d}h(y)\rho(y)\,\mathrm{d}y=0.$$

在这种情况下, (13.2.7) 的解在相差一个常数的前提下是唯一的. 当 $h = b$ 时, 由偏微分方程 (13.2.1) 产生向量场, 上述条件变为

$$\int_{\mathbb{T}^d} b(y)\rho(y)\,\mathrm{d}y = 0, \tag{13.2.8}$$

称其为**中心条件**. 通过令 (13.2.7) 的解满足

$$\int_{\mathbb{T}^d} v(y)\rho(y)\,\mathrm{d}y = 0 \tag{13.2.9}$$

来固定解 (13.2.7) 中的自由常数.

当不满足中心条件时, 必须用不同的方式来重新标定初始问题以考察有效的平流行为. 特别地, (13.2.3) 可被**平流尺度**代替

$$x \to \varepsilon^{-1}x, \quad t \to \varepsilon^{-1}t. \tag{13.2.10}$$

然后使用平均化来找到有效方程, 此时该方程是一个运输方程, 见第 14 章.

13.3 简化方程

假设向量场 $b(y)$ 满足中心条件 (13.2.8). 定义向量场 $\chi(y)$ 为下列单胞问题的解:

$$-\mathcal{L}_0\chi = b, \quad \chi\text{的周期为 } 1, \quad \int_{\mathbb{T}^d} \chi(y)\rho(y)\,\mathrm{d}y = 0. \tag{13.3.1}$$

有效扩散张量 (或者是有效扩散系数) 定义如下:

$$\mathcal{K} = DI + 2D\int_{\mathbb{T}^d} \nabla_y\chi(y)^{\mathrm{T}}\rho(y)\,\mathrm{d}y + \int_{\mathbb{T}^d} (b(y)\otimes\chi(y))\rho(y)\,\mathrm{d}y. \tag{13.3.2}$$

结论 13.1 假设 (13.2.8) 成立, 对于 $0 < \varepsilon \ll 1$ 和 $O(1)$ 的时间 t, (13.2.4) 的解 u^ε 可由 u 进行逼近, u 是下面均匀化方程的解:

$$\frac{\partial u}{\partial t} = \mathcal{K} : \nabla_x\nabla_x u, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}^+, \tag{13.3.3a}$$

$$u = g, \quad (x,t) \in \mathbb{R}^d \times \{0\}. \tag{13.3.3b}$$

注 13.2 既然 Hessian 矩阵 $\nabla_x\nabla_x u$ 具有对称性, 则将性质 (2.2.2) 应用到 (13.3.2) 上可得下面关于 $\mathcal{K}$ 的表达式:

$$\mathcal{K} = DI + D\int_{\mathbb{T}^d} \left(\nabla_y\chi(y) + \nabla_y\chi(y)^{\mathrm{T}}\right)\rho(y)\,\mathrm{d}y + \frac{1}{2}\int_{\mathbb{T}^d} (b(y)\otimes\chi(y) + \chi(y)\otimes b(y))\,\rho(y)\,\mathrm{d}y. \tag{13.3.4}$$

应用这种想法可能会出现很多变量. □

13.4 推 导

现在的目标就是使用多尺度方法来分析 $u^\varepsilon(x,t)$ 的行为, 其中当 $\varepsilon\to 0$ 时, $u^\varepsilon(x,t)$ 为 (13.2.4) 的解. 特别地, 想推得结论 13.1.

引入辅助变量[①] $y=x/\varepsilon$. 设 $\phi=\phi(x,x/\varepsilon)$ 为标量值. 由链式法则给出

$$\nabla\phi=\nabla_x\phi+\frac{1}{\varepsilon}\nabla_y\phi,\quad \Delta\phi=\Delta_x\phi+\frac{2}{\varepsilon}\nabla_x\cdot\nabla_y\phi+\frac{1}{\varepsilon^2}\Delta_y\phi.$$

方程 (13.2.4) 右边出现的偏微分算子变为

$$\mathcal{L}=\frac{1}{\varepsilon^2}\mathcal{L}_0+\frac{1}{\varepsilon}\mathcal{L}_1+\mathcal{L}_2,$$

其中

$$\begin{aligned}\mathcal{L}_0&=b(y)\cdot\nabla_y+D\Delta_y,\\ \mathcal{L}_1&=b(y)\cdot\nabla_x+2D\nabla_x\cdot\nabla_y,\\ \mathcal{L}_2&=D\Delta_x.\end{aligned}$$

相对于 x 和 y, 方程 (13.2.4a) 变为

$$\frac{\partial u^\varepsilon}{\partial t}=\left(\frac{1}{\varepsilon^2}\mathcal{L}_0+\frac{1}{\varepsilon}\mathcal{L}_1+\mathcal{L}_2\right)u^\varepsilon.$$

希望找到一个形如下面多尺度展开式的解:

$$u^\varepsilon(x,t)=u_0(x,y,t)+\varepsilon u_1(x,y,t)+\varepsilon^2u_2(x,y,t)+\cdots,\tag{13.4.1}$$

其中 $u_j(x,y,t)\,(j=1,2,\cdots)$ 在周期为 1 的 y 上具有周期性. 将其代入 (13.4.1) 中并考虑 ε 上等次幂系数, 得到下面的方程:

$$O\left(\frac{1}{\varepsilon^2}\right):-\mathcal{L}_0u_0=0;\tag{13.4.2a}$$

$$O\left(\frac{1}{\varepsilon}\right):-\mathcal{L}_0u_1=\mathcal{L}_1u_0;\tag{13.4.2b}$$

$$O(1):-\mathcal{L}_0u_2=\mathcal{L}_1u_1+\mathcal{L}_2u_0-\frac{\partial u_0}{\partial t}.\tag{13.4.2c}$$

注意到仅在 y 上是一个微分算子的 $\mathcal{L}_0$ 具有周期边界条件.

因为 $\mathcal{L}_0$ 具有一个一维零空间, 方程 (13.4.2a) 暗含表达式的第一项独立于 y, 使得只有 $u_0=u(x,t)$. 注意到

$$\mathcal{L}_1u_0=b(y)\cdot\nabla_xu(x,t).$$

① 在椭圆情形下, 这就是使用了尺度分离的假设, 将 x 和 y 视为独立变量. 当 $\varepsilon\to 0$ 时, 证明这个假设是均匀化理论中重要的问题之一, 见第 20 章.

通过 Fredholm 性质, 中心条件 (13.2.8) 确保 (13.4.2b) 有解. 因为 $\mathcal{L}_0$ 仅在 y 上是一个微分算子, 可以使用变量分离将解记为

$$u_1(x,y,t) = \chi(y) \cdot \nabla_x u(x,t),$$

则 $\chi(y)$ 为单胞问题 (13.3.1) 的解. 所作的假设暗含对于单胞问题这里存在一个唯一的光滑解.

现在继续分析 $O(1)$ 方程 (13.4.2c). 可解性条件 (13.2.8) 可理解为

$$\int_{\mathbb{T}^d} \left(\frac{\partial u_0}{\partial t} - \mathcal{L}_2 u_0 - \mathcal{L}_1 u_1 \right) \rho \mathrm{d}y = 0.$$

由 $u_0 = u(x,t)$ 独立于 y 将前面的方程记为下面的形式:

$$\frac{\partial u}{\partial t} = D\Delta u + \int_{\mathbb{T}^d} (\mathcal{L}_1 u_1) \rho \mathrm{d}y. \tag{13.4.3}$$

现在有

$$\mathcal{L}_1 u_1 = (b \cdot \nabla_x (\chi \cdot \nabla_x u) + 2D\nabla_x \cdot \nabla_y (\chi \cdot \nabla_x u)) = \left(b \otimes \chi + 2D\nabla_y \chi^{\mathrm{T}}\right) : \nabla_x \nabla_x u.$$

根据前面的计算结果, 方程 (13.4.3) 变为

$$\frac{\partial u}{\partial t} = \mathcal{K} : \nabla_x \nabla_x u,$$

该方程为均匀化方程 (13.3.3a). 有效扩散系数 $\mathcal{K}$ 由式 (13.3.2) 计算得到.

13.5 简化方程的性质

本节中将得出有效扩散系数是一致正定的, 这暗含均匀化方程具有适定性. 为了证明该结论, 需要计算关于算子 $\mathcal{L}_0$ 的 Dirichlet 形式. 下面是对附加噪声使用定理 6.12 得到的直接结果. 回想 ρ 是不变分布, 在零空间 $\mathcal{L}_0^*$ 上是一个非负 $L^1(\mathbb{T}^d)$ 函数.

引理 13.3 设 $f(y) \in C^2_{\mathrm{per}}(\mathbb{T}^d)$, 则

$$\int_{\mathbb{T}^d} (-\mathcal{L}_0 f(y)) f(y) \rho(y) \mathrm{d}y = D \int_{\mathbb{T}^d} |\nabla_y f(y)|^2 \rho(y) \mathrm{d}y. \tag{13.5.1}$$

注 13.4 设 $L^2_\rho(\mathbb{T}^d)$ 为不变分布 $\rho(y)$ 加权的 L^2 空间, 并且分别用 $(\cdot,\cdot)_{L^2_\rho}$ 和 $\|\cdot\|_{L^2_\rho}$ 定义内积和相应的范数, 则由方程 (6.3.13), 可将引理 13.3 的结论表示为下面的形式:

$$(-\mathcal{L}_0 f, f)_{L^2_\rho} = D \|\nabla_y f\|^2_{L^2_\rho}. \qquad \square$$

本节主要的结论是有效扩散系数是一个一致正定矩阵. 特别地, 有如下定理:

定理 13.5　设 $\xi \in \mathbb{R}^d$ 是一个任意向量, 并且设 $\chi_\xi(y) := \chi(y) \cdot \xi$, 则

$$\langle \xi, \mathcal{K}\xi \rangle = D \int_{\mathbb{T}^d} |\xi + \nabla_y \chi_\xi(y)|^2 \rho(y) \, \mathrm{d}y,$$

而且

$$\alpha |\xi|^2 \leqslant \langle \xi, \mathcal{K}\xi \rangle, \quad \forall \xi \in \mathbb{R}^d, \tag{13.5.2}$$

其中

$$\alpha = D \left(\int_{\mathbb{T}^d} \rho^{-1}(y) \, \mathrm{d}y \right)^{-1}. \tag{13.5.3}$$

证明　注意到 $-\mathcal{L}_0 \chi_\xi = \xi \cdot b$. 由 $\mathcal{K}$ 的定义以及引理 13.3 计算得

$$\begin{aligned}
\langle \xi, \mathcal{K}\xi \rangle &= D|\xi|^2 + 2D \int_{\mathbb{T}^d} \xi \cdot \nabla_y \chi_\xi(y) \rho(y) \, \mathrm{d}y + \int_{\mathbb{T}^d} (\xi \cdot b) \chi_\xi(y) \rho(y) \, \mathrm{d}y \\
&= D|\xi|^2 + 2D \int_{\mathbb{T}^d} \xi \cdot \nabla_y \chi_\xi(y) \rho(y) \, \mathrm{d}y + D \int_{\mathbb{T}^d} |\nabla_y \chi_\xi(y)|^2 \rho(y) \, \mathrm{d}y \\
&= D \int_{\mathbb{T}^d} |\xi + \nabla_y \chi_\xi(y)|^2 \rho(y) \, \mathrm{d}y.
\end{aligned}$$

事实上, 由前面的方程立即可得有效扩散系数是非负定的. 为了表明 $\mathcal{K}$ 是一致正定的, 应用在 $\mathbb{T}^d$ 上周期函数导数的积分为 0, Cauchy-Schwarz 不等式以及 $\rho(y)$ 处处为正等事实来进行计算,

$$\begin{aligned}
D|\xi|^2 &= D \left| \int_{\mathbb{T}^d} (\xi + \nabla_y \chi_\xi) \, \mathrm{d}y \right|^2 \\
&= D \left| \int_{\mathbb{T}^d} (\xi + \nabla_y \chi_\xi) \rho^{1/2}(y) \rho^{-1/2}(y) \, \mathrm{d}y \right|^2 \\
&\leqslant D \left(\int_{\mathbb{T}^d} |\xi + \nabla_y \chi_\xi|^2 \rho(y) \, \mathrm{d}y \right) \left(\int_{\mathbb{T}^d} \rho^{-1}(y) \, \mathrm{d}y \right) \\
&= \langle \xi, \mathcal{K}\xi \rangle \left(\int_{\mathbb{T}^d} \rho^{-1}(y) \, \mathrm{d}y \right),
\end{aligned}$$

由此立即可得到下界. □

弄明白如何将有效扩散张量 $\mathcal{K}$ 与初始扩散张量 DI 进行比较是件有趣的事情. 其证明了 $\mathcal{K}$ 较 D 可能更大, 也可能更小 (就矩阵而言). 这个问题将在下节中进行详细的讨论, 在那里将证明对于梯度向量场 b, 有效扩散系数比 D 更小; 对于无散向量场 b, 有效扩散系数比 D 更大.

13.6 应　用

本节对 (13.2.4a) 中漂移项 b 的两种特殊情况 (梯度场和无散场) 进行讨论. 在这两种情形下, 可以进行显式计算, 这样可以得到一些重要的结论. 特别地, 将得到一个 (唯一) 不变分布公式, 并且通过它来简化中心条件 (13.2.8), 还可以将有效扩散系数与初始扩散系数 D 进行比较. 将会看到相对于梯度向量场 b, 有效扩散系数较 D 更小; 相对于无散向量场 b, 有效扩散系数较 D 更大. 还将研究梯度流和无散流这两种特殊情况, 推出了有效扩散系数的近似公式.

为什么说思考梯度流和无散流这两种情况是有趣的事, 这里至少有两方面的原因. 一方面, 形如 (13.2.1) 的椭圆型偏微分方程常在应用中出现, 其中 b 为标量场的梯度或者是无散场. 当 $b = -\nabla V$ 时, 方程 (13.2.1) 描述的是在周期势下的 Brownian 运动. 另一方面, 当 b 是无散场时, 方程 (13.2.1) 变为对流扩散方程, 该方程描述了不可压缩流体中的混合过程. 根据 Hodge 分解理论, $\mathbb{T}^d$ 上的每个光滑向量场都能被分解成梯度场与无散场之和,

$$b(y) = -\nabla V(y) + v(y), \quad \nabla \cdot v(y) = 0,$$

其中

$$(-\nabla V(y), v(y))_{L^2(\mathbb{T}^d)} = 0,$$

从而通过研究梯度流和无散流来研究这个分解的两个极端情形.

13.6.1 梯度向量场

考虑如下情形, 即方程 (13.2.4a) 中向量场 $b(y)$ 是一个光滑标量周期函数的梯度:

$$b(y) = -\nabla_y V(y), \tag{13.6.1}$$

称函数 V 为势能. 在这种情况下, 可直接推出稳态伴随方程 (13.2.6) 解 ρ 的计算公式, 方程 (13.2.6) 具有周期边界条件.

引理 13.6 假设向量场 b 是一个由 (13.6.1) 给出的梯度. 设 $\mathcal{L}_0^*$ 为 $\mathcal{L}_0$ 的伴随, $\mathcal{L}_0$ 的定义见 (13.2.5), 则方程

$$\mathcal{L}_0^*\rho = 0, \quad \int_{\mathbb{T}^d} \rho(y)\mathrm{d}y = 1 \tag{13.6.2}$$

在 $\mathbb{T}^d$ 上服从周期边界条件, 并且方程有唯一解, 形式如下:

$$\rho(y) = \frac{1}{Z}\mathrm{e}^{-V(y)/D}, \quad Z = \int_{\mathbb{T}^d} \mathrm{e}^{-V(y)/D}\mathrm{d}y. \tag{13.6.3}$$

证明　根据方程 (13.6.1), 方程 (13.6.2) 变为

$$\nabla_y \cdot (\nabla_y V(y)\rho(y) + D\nabla_y \rho(y)) = 0. \tag{13.6.4}$$

立即可验证由 (13.6.3) 给出的 $\rho(y)$ 满足

$$\nabla_y V(y)\rho(y) + D\nabla_y \rho(y) = 0,$$

因此, 也满足 (13.6.4), 而且通过构造可得

$$\int_{\mathbb{T}^d} \frac{1}{Z} \mathrm{e}^{-V(y)/D} \mathrm{d}y = 1,$$

所以 $\rho(y)$ 可以规范化. 这样就已经构造了方程 (13.6.2) 的解, 由生成子为 $\mathcal{L}_0$ 的随机过程的遍历性可得解是唯一的 (见定理 6.16). □

注 13.7　称 (13.2.8) 定义的正函数 ρ 为**Gibbs 分布**, 并称概率测度 $\rho(y)\,\mathrm{d}y$ 为**Gibbs 测度**, 称归一化常数 Z 为**配分函数**. □

对任意势能, 梯度流恒满足中心条件 (13.2.8).

引理 13.8　考虑由 (13.2.5) 给出的具有周期边界条件的算子 $\mathcal{L}_0$, 并假设 $b(y) = -\nabla_y V(y)$, 其中 $V \in C^1_{\mathrm{per}}(\mathbb{T}^d)$, 则总可以满足中心条件 (13.2.8).

证明　使用收敛定理计算

$$\begin{aligned}\int_{\mathbb{T}^d} b(y)\rho(y)\,\mathrm{d}y &= \frac{1}{Z}\int_{\mathbb{T}^d} -\nabla_y V(y)\,\mathrm{e}^{-V(y)/D}\mathrm{d}y \\ &= \frac{D}{Z}\int_{\mathbb{T}^d} \nabla_y \mathrm{e}^{-V(y)/D}\mathrm{d}y \\ &= 0.\end{aligned}$$

□

对于梯度流来说, (13.2.5) 中定义的具有周期边界条件的算子 $\mathcal{L}_0$ 在合适的函数空间上变得具有对称性, 有如下定理:

引理 13.9　假设满足条件 (13.6.1), 并且令 ρ 定义为 Gibbs 分布 (13.6.3), 则对于所有 $f, h \in C^2_{\mathrm{per}}(\mathbb{T}^d)$, (13.2.5) 中给出的算子 $\mathcal{L}_0$ 满足

$$\int_{\mathbb{T}^d} f(y)(\mathcal{L}_0 h(y))\rho(y)\,\mathrm{d}y = \int_{\mathbb{T}^d} h(y)(\mathcal{L}_0 f(y))\rho(y)\,\mathrm{d}y. \tag{13.6.5}$$

证明　使用收敛定理有

$$\begin{aligned}\int_{\mathbb{T}^d} f\mathcal{L}_0 h\rho\mathrm{d}y &= \frac{1}{Z}\int_{\mathbb{T}^d} f(-\nabla_y V\cdot\nabla_y h)\,\mathrm{e}^{-V/D}\mathrm{d}y + \frac{D}{Z}\int_{\mathbb{T}^d} f\Delta_y h\mathrm{e}^{-V/D}\mathrm{d}y \\ &= \frac{D}{Z}\int_{\mathbb{T}^d} f\nabla_y h\cdot\nabla_y\left(\mathrm{e}^{-V/D}\right)\mathrm{d}y - \frac{D}{Z}\int_{\mathbb{T}^d}(\nabla_y f\cdot\nabla_y h)\,\mathrm{e}^{-V/D}\mathrm{d}y\end{aligned}$$

$$-\frac{D}{Z}\int_{\mathbb{T}^d} f\nabla_y h\cdot\nabla_y\left(\mathrm{e}^{-V/D}\right)\mathrm{d}y=-D\int_{\mathbb{T}^d}\left(\nabla_y f\cdot\nabla_y h\right)\rho\mathrm{d}y.$$

由于上述表达式的最后一行在 f,h 上是对称的, 因而有 (13.6.5). □

注 13.10 由证明定理 6.12 时用到的恒等式 (6.3.11) 很自然地得到 $\mathcal{L}_0$ 是对称的, 而且由证明引理 13.9 时用到的计算得到如下公式:

$$\int_{\mathbb{T}^d} f\left(-\mathcal{L}_0 h\right)\rho\mathrm{d}y=D\int_{\mathbb{T}^d}\left(\nabla_y f\cdot\nabla_y h\right)\rho\mathrm{d}y \tag{13.6.6}$$

对于所有 $f,h\in C^2_{\mathrm{per}}\left(\mathbb{T}^d\right)$. 当 $f=h$ 时, 从式 (13.6.6) 可得到 Dirichlet 形式 (引理 13.3). 现在设 $\phi,\psi\in C^2_{\mathrm{per}}\left(\mathbb{T}^d;\mathbb{R}^d\right)$, 则根据 (13.6.6) 有

$$\int_{\mathbb{T}^d}\left(\phi\otimes\left(-\mathcal{L}_0\psi\right)\right)\rho\mathrm{d}y=D\int_{\mathbb{T}^d}\left(\nabla_y\phi\otimes\nabla_y\psi\right)\rho\mathrm{d}y. \qquad\square \tag{13.6.7}$$

注 13.11 使用注 13.4 中引入的定义, 通过将 $\mathcal{L}_0$ 看成是从 L^2_ρ 到 L^2_ρ 上具有对称性的算子可表示引理 13.9 的结论, 而且式 (13.6.6) 可记为如下形式:

$$\left(f,-\mathcal{L}_0 h\right)_{L^2_\rho}=D\left(\nabla_y f,\nabla_y h\right)_{L^2_\rho}.$$

称由空间 L^2_ρ 上的对称生成子生成的遍历性 Markov 过程是**可逆的**. 这就说明含有附加噪声并且漂移项是一个梯度场的随机微分方程系统是可逆的. □

现在准备证明有效扩散系数的不同性质. 为此, 需要下面的分部积分公式, 该公式是由收敛定理以及 χ 和 ρ 的周期性得出的, 公式如下:

$$\int_{\mathbb{T}^d}\left(\nabla_y\chi\right)\rho\mathrm{d}y=\int_{\mathbb{T}^d}\left(\nabla_y\left(\chi\rho\right)-\chi\otimes\nabla_y\rho\right)\mathrm{d}y=-\int_{\mathbb{T}^d}\left(\chi\otimes\nabla_y\rho\right)\mathrm{d}y. \tag{13.6.8}$$

定理 13.12 假设 $b(y)$ 是一个梯度, 使得 (13.6.1) 成立, 令 $\rho(y)$ 定义为 Gibbs 分布 (13.6.3), 则有效扩散系数 (13.3.2) 满足上、下界:

$$\frac{D}{Z\hat{Z}}\leqslant\langle\xi,\mathcal{K}\xi\rangle\leqslant D\left|\xi\right|^2,\quad\forall\xi\in\mathbb{R}^d, \tag{13.6.9}$$

其中

$$\hat{Z}=\int_{\mathbb{T}^d}\mathrm{e}^{V(y)/D}\mathrm{d}y.$$

特别地, 与分子扩散系数相比, 扩散总是亏损的, 而且有效扩散系数具有对称性①.

① 注意到 Cauchy-Schwarz 不等式表明 $Z\hat{Z}\geqslant 1$.

证明　由一般下界 (13.5.2), 方程 (13.5.3) 以及 Gibbs 测度公式得到下界. 为了确定上界, 使用 (13.6.8) 和 (13.6.7) 得

$$
\begin{aligned}
\mathcal{K} &= DI + 2D\int_{\mathbb{T}^d} (\nabla\chi)^{\mathrm{T}}\rho \mathrm{d}y + \int_{\mathbb{T}^d} -\nabla_y V\otimes\chi\rho\mathrm{d}y \\
&= DI - 2D\int_{\mathbb{T}^d} \nabla_y\rho\otimes\chi\mathrm{d}y + \int_{\mathbb{T}^d} -\nabla_y V\otimes\chi\rho\mathrm{d}y \\
&= DI - 2\int_{\mathbb{T}^d} -\nabla_y V\otimes\chi\rho\mathrm{d}y + \int_{\mathbb{T}^d} -\nabla_y V\otimes\chi\rho\mathrm{d}y \\
&= DI - \int_{\mathbb{T}^d} -\nabla_y V\otimes\chi\rho\mathrm{d}y \\
&= DI - \int_{\mathbb{T}^d} (-\mathcal{L}_0\chi)\otimes\chi\rho\mathrm{d}y \\
&= DI - D\int_{\mathbb{T}^d} (\nabla_y\chi\otimes\nabla_y\chi)\,\rho\mathrm{d}y.
\end{aligned}
\tag{13.6.10}
$$

因此, 对于 $\chi_\xi = \chi\cdot\xi$,

$$
\langle\xi,\mathcal{K}\xi\rangle = D\,|\xi|^2 - D\int_{\mathbb{T}^d} |\nabla_y\chi_\xi|^2\,\rho\mathrm{d}y \leqslant D\,|\xi|^2 .
$$

这就证明了亏损. 由 (13.6.10) 可知 $\mathcal{K}$ 具有对称性. □

一维例子

一维的例子总有梯度形式 $b(y) = -\partial_y V(y)$, 而且在一维情形下可以以封闭形式来求解单胞问题 (13.3.1), 并且计算有效扩散系数的显式形式. 首先计算扩散系数的结构,

$$
\begin{aligned}
\mathcal{K} &= D + 2D\int_0^1 \partial_y\chi\rho\mathrm{d}y + \int_0^1 -\partial_y V\chi\rho\mathrm{d}y \\
&= D + 2D\int_0^1 \partial_y\chi\rho\mathrm{d}y + D\int_0^1 \chi\partial_y\rho\mathrm{d}y \\
&= D + 2D\int_0^1 \partial_y\chi\rho\mathrm{d}y - D\int_0^1 \partial_y\chi\rho\mathrm{d}y \\
&= D\int_0^1 (1+\partial_y\chi)\,\rho\mathrm{d}y.
\end{aligned}
\tag{13.6.11}
$$

一维情形下的单胞问题为

$$
D\partial_{yy}\chi - \partial_y V\partial_y\chi = \partial_y V. \tag{13.6.12}
$$

用 $\mathrm{e}^{-V(y)/D}$ 右乘方程 (13.6.12) 得

$$
\partial_y\left(\partial_y\chi\mathrm{e}^{-V(y)/D}\right) = -\partial_y\left(\mathrm{e}^{-V(y)/D}\right).
$$

将此方程从 0 到 y 上积分, 并且乘以 $\mathrm{e}^{V(y)/D}$ 得

$$\partial_y \chi(y) = -1 + c_1 \mathrm{e}^{V(y)/D},$$

再次积分可得

$$\chi(y) = -y + c_1 \int_0^y \mathrm{e}^{V(y)/D} \mathrm{d}y + c_2.$$

周期边界条件暗含了 $\chi(0) = \chi(1)$, 由此可以推得

$$-1 + c_1 \int_0^1 \mathrm{e}^{V(y)/D} \mathrm{d}y = 0.$$

因此,

$$c_1 = \frac{1}{\hat{Z}}, \quad \hat{Z} = \int_0^1 \mathrm{e}^{V(y)/D} \mathrm{d}y,$$

推得

$$\partial_y \chi = -1 + \frac{1}{\hat{Z}} \mathrm{e}^{V(y)/D}.$$

将此表达式代入 (13.6.11) 得

$$\mathcal{K} = \frac{D}{Z} \int_0^1 (1 + \partial_y \chi(y)) \mathrm{e}^{-V(y)/D} \mathrm{d}y = \frac{D}{Z\hat{Z}} \int_0^1 \mathrm{e}^{V(y)/D} \mathrm{e}^{-V(y)/D} \mathrm{d}y = \frac{D}{Z\hat{Z}}, \tag{13.6.13}$$

其中

$$Z = \int_0^1 \mathrm{e}^{-V(y)/D} \mathrm{d}y, \quad \hat{Z} = \int_0^1 \mathrm{e}^{V(y)/D} \mathrm{d}y. \tag{13.6.14}$$

注意到在一维情形下有效系数公式严格为 (13.6.9) 中的下界, 这表明下界是可以达到的.

例 13.13　考虑势

$$V(y) = \begin{cases} a_1, & y \in \left[0, \dfrac{1}{2}\right], \\ a_2, & y \in \left(\dfrac{1}{2}, 1\right], \end{cases} \tag{13.6.15}$$

其中 a_1 和 a_2 均为正常数①.

直接计算 (13.6.14) 中的积分得如下公式:

$$\mathcal{K} = \frac{D}{\cosh^2\left(\dfrac{a_1 - a_2}{D}\right)}. \tag{13.6.16}$$

① 当然, 这个势是不连续的, 更不用说是光滑的, 因而不能应用本章中发展的理论. 然而, 可以考虑这种不连续势的规则化形式, 那样就可以使用均匀化理论.

图 13.1 绘制出由 (13.6.16) 获取的扩散系数, 它是基于对数尺度下分子扩散系数 D 的一个函数. 注意到当极限 $D \to 0$ 时, $\mathcal{K}$ 呈指数衰减. □

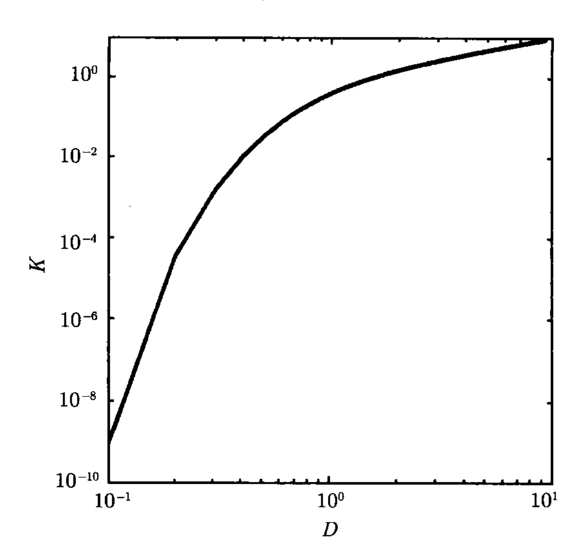


图 13.1　基于势 (13.6.15) 有效扩散率对分子扩散率的对数图

13.6.2　无散场

本节考虑 (13.2.4a) 的均匀化问题, 此时向量场 $b(y)$ 是无散的 (或是不可压缩的),

$$\nabla \cdot b(y) = 0. \tag{13.6.17}$$

因为具有周期边界条件的对流算子

$$\hat{\mathcal{L}}_0 = b(y) \cdot \nabla_y$$

在 $L^2(\mathbb{T}^d)$ 上是具有反对称性的, 所以 $b(y)$ 的不可压缩性就大大简化了分析.

引理 13.14　设 $b(y) \in C^1_{\text{per}}(\mathbb{T}^d; \mathbb{R}^d)$ 满足 (13.6.17), 则对于所有的 $f(y), h(y) \in C^1_{\text{per}}(\mathbb{T}^d)$ 有

$$\int_{\mathbb{T}^d} f(y)(b(y) \cdot \nabla_y h(y))\,\mathrm{d}y = -\int_{\mathbb{T}^d} h(y)(b(y) \cdot \nabla_y f(y))\,\mathrm{d}y.$$

特别地,

$$\int_{\mathbb{T}^d} f(y)(b(y) \cdot \nabla_y f(y))\,\mathrm{d}y = 0. \tag{13.6.18}$$

证明　由 $b(y)$ 的不可压缩性以及 $f(y), h(y)$ 和 $b(y)$ 的周期性计算

$$\begin{aligned}\int_{\mathbb{T}^d} f(y)(b(y)\cdot\nabla_y h(y))\,\mathrm{d}y &= \int_{\mathbb{T}^d} f(y)\nabla_y\cdot(b(y)h(y))\,\mathrm{d}y \\ &= -\int_{\mathbb{T}^d} \nabla_y f(y)\cdot(b(y)h(y))\,\mathrm{d}y \\ &= -\int_{\mathbb{T}^d} h(y)(b(y)\cdot\nabla_y f(y))\,\mathrm{d}y.\end{aligned}$$

在该计算中, 当 $f=h$ 时可得方程 (13.6.18). □

使用前面的引理很容易证明快速过程的唯一不变测度是 Lebesgue 测度.

引理 13.15　令 $\mathcal{L}_0$ 为 (13.2.5) 中定义的具有周期边界条件的算子, 并且 $b(y)$ 满足 (13.6.17). 令 $\mathcal{L}_0^*$ 定义为 $\mathcal{L}_0$ 的 L^2 伴随, 则在 $\mathbb{T}^d$ 上具有周期边界条件的伴随方程

$$\mathcal{L}_0^*\rho=0,\quad \int_{\mathbb{T}^d}\rho(y)\,\mathrm{d}y=1 \tag{13.6.19}$$

有唯一古典解, 该解由下式给出:

$$\rho(y)=1. \tag{13.6.20}$$

证明　引理 13.14 暗含 $\mathcal{L}_0$ 的 L^2 伴随为

$$\mathcal{L}_0^*=-b(y)\cdot\nabla_y+D\Delta_y, \tag{13.6.21}$$

其具有周期边界条件. 设 $\rho(y)$ 为方程 (13.6.19) 的一个解, 用 $\rho(y)$ 乘以方程, 并在 $\mathbb{T}^d$ 上积分, 由引理 13.14 得

$$\int_{\mathbb{T}^d}|\nabla_y\rho(y)|^2\mathrm{d}y=0, \tag{13.6.22}$$

由方程 (13.6.22) 推得 $\rho(y)$ 是一个常数. 因此, 由 (13.6.20) 得到 (13.6.19) 的唯一标准解. □

注 13.16　通过观察, 可将解 $\rho(y)=1$ 看成是在 (13.6.21) 的零空间上, 则由生成子为 $\mathcal{L}_0$ 的过程的遍历性 (见定理 6.16) 或最大值定理可以证明解的唯一性. □

注 13.17　由命题 13.15 可立即得到一个结论, 即对于无散场, 可解性条件 (13.2.8) 变为

$$\int_{\mathbb{T}^d}b(y)\mathrm{d}y=0.$$

因此, 可以直接考察一个给定的周期性无散场是否满足可解性条件 —— 场必须在单位圆环上均值为零. □

现在设 $\chi(y)$ 为单胞问题 (13.3.1) 的解, 其中 $b(y)$ 满足 (13.6.17). $\chi(y)$ 的周期性以及 (13.6.20) 暗含了方程 (13.3.2) 右边的第二项消失, 并且有效扩散系数方程降为

$$\mathcal{K} = DI + \int_{\mathbb{T}^d} b(y) \otimes \chi(y) \, \mathrm{d}y. \tag{13.6.23}$$

对于梯度流, (13.3.2) 给出的有效扩散系数是对称的. 对于无散场来说, 这是不可能的. 但是由注 13.2 可知, 仅有 $\mathcal{K}$ 的对称部分出现在均匀化方程中. 因此, 重新将有效扩散系数定义为 $\mathcal{K}$ 的对称部分,

$$\mathcal{K} := DI + \frac{1}{2} \int_{\mathbb{T}^d} (b(y) \otimes \chi(y) + \chi(y) \otimes b(y)) \, \mathrm{d}y. \tag{13.6.24}$$

目的就是表明均匀化过程增强了扩散性, 即有效扩散性总是较分子扩散性 D 更强. 为此, 需要一个 $\mathcal{K}$ 的替换表达式.

定理 13.18 由表达式 (13.6.24) 给出的有效扩散系数 $\mathcal{K}$ 可记为如下形式:

$$\mathcal{K} = DI + D \int_{\mathbb{T}^d} \nabla_y \chi(y) \otimes \nabla_y \chi(y) \, \mathrm{d}y. \tag{13.6.25}$$

证明 取单胞问题的外积, 其中 $\chi(y)$ 在单胞上积分时单胞区域在左边,

$$-D \int_{\mathbb{T}^d} \chi(y) \otimes \Delta_y \chi(y) \, \mathrm{d}y - \int_{\mathbb{T}^d} \chi(y) \otimes (\nabla_y \chi(y) \, b(y)) \, \mathrm{d}y = \int_{\mathbb{T}^d} \chi(y) \otimes b(y) \, \mathrm{d}y.$$

在方程左边的两个积分上应用收敛定理, 使用遍历性以及 b 是无散场可得

$$D \int_{\mathbb{T}^d} \nabla_y \chi(y) \otimes \nabla_y \chi(y) \, \mathrm{d}y + \int_{\mathbb{T}^d} (\nabla \chi(y) \, b(y)) \otimes \chi(y) \, \mathrm{d}y = \int_{\mathbb{T}^d} \chi(y) \otimes b(y) \, \mathrm{d}y, \tag{13.6.26}$$

或者也可以对方程 (13.3.1) 的右边取外积, 仅在第一个积分上使用收敛定理得

$$D \int_{\mathbb{T}^d} \nabla_y \chi(y) \otimes \nabla_y \chi(y) \, \mathrm{d}y - \int_{\mathbb{T}^d} (\nabla_y \chi(y) \, b(y)) \otimes \chi(y) \, \mathrm{d}y = \int_{\mathbb{T}^d} b(y) \otimes \chi(y) \, \mathrm{d}y. \tag{13.6.27}$$

将方程 (13.6.26) 与方程 (13.6.27) 相加得

$$\frac{1}{2} \int_{\mathbb{T}^d} (b(y) \otimes \chi(y) + \chi(y) \otimes b(y)) \, \mathrm{d}y = D \int_{\mathbb{T}^d} \nabla_y \chi(y) \otimes \nabla_y \chi(y) \, \mathrm{d}y.$$

将这个方程代入方程 (13.6.24) 得到方程 (13.6.25). □

现在就能得到有效扩散系数的上、下界了.

定理 13.19 假设 $b(y)$ 是无散的, 则有效扩散系数满足上、下界:

$$D |\xi|^2 \leqslant \langle \xi, \mathcal{K}\xi \rangle \leqslant \left(D + \frac{C}{D} \right) |\xi|^2, \tag{13.6.28}$$

其中 $C=C(b,\Omega)>0$ 是可以显式计算的①. 对于所有的 ξ, 仅当 $b(y)\equiv 0$ 时, 下界变为一个等式.

证明　由一般边界 (13.5.2), 方程 (13.5.3) 以及对于无散流 $\rho(y)=1$ 这一事实可得下界, 而且方程 (13.6.25) 暗含

$$\langle\xi,\mathcal{K}\xi\rangle := D|\xi|^2 + D\int_{\mathbb{T}^d}|\nabla_y\chi_\xi(y)|^2\mathrm{d}y, \tag{13.6.29}$$

其中 $\chi_\xi=\chi\cdot\xi$. 显然, 对于所有的 ξ, 等式 $\langle\xi,\mathcal{K}\xi\rangle=D|\xi|^2$ 暗含 $\chi_\xi=0$, 并且对于所有 ξ 暗含 $\chi(y)\equiv 0$. 由 (13.6.1) 可知这暗含了 $b\equiv 0$.

对于上界, 将单胞问题与任意向量场 $\xi\in\mathbb{R}^d$ 作内积得

$$-\mathcal{L}_0\chi_\xi=b\cdot\xi.$$

将该方程乘以 χ_ξ 并在 $\mathbb{T}^d$ 上积分, 使用引理 13.14 和 Poincaré不等式计算

$$D\left\|\nabla_y\chi^\xi\right\|_{L^2}^2=(-\mathcal{L}_0\chi_\xi,\chi_\xi)=(b\cdot\xi,\chi_\xi)\leqslant\|b\cdot\xi\|_{L^2}\|\chi_\xi\|_{L^2}\leqslant C_p\|b\|_{L^2}\|\nabla_y\chi_\xi\|_{L^2}|\xi|,$$

其中 C_p 是 $\mathbb{T}^d$ 上的 Poincaré常数. 从这个估计值可以推得

$$\|\nabla_y\chi_\xi\|_{L^2}\leqslant\frac{\sqrt{C}}{D}|\xi|,$$

其中 $C=(C_p\|b\|_{L^2})^2$. 由 (13.6.29) 可得出证明的结论. □

二维剪切流

本节将研究一个无散流 (一个剪切流) 的例子. 对于该无散流, 单胞问题有闭合形式的解. 剪切速度场的结构使得单胞问题变成一个常微分方程.

设 $y=(y_1,y_2)^{\mathrm{T}}$, 在下面二维情形的速度场:

$$b(y)=(0,b_2(y_1))^{\mathrm{T}} \tag{13.6.30}$$

中考虑 (13.2.4a) 的均匀化问题, 其中 $b_2(y_1)$ 是光滑的、周期为 1 且均值为零的函数. 注意到速度场 (13.6.30) 是不可压缩的,

$$\nabla\cdot b(y)=\frac{\partial b_1}{\partial y_1}+\frac{\partial b_2}{\partial y_2}=\frac{\partial b_2(y_1)}{\partial y_2}=0.$$

单胞问题的两部分满足

$$-D\Delta_y\chi_1(y)-b_2(y_1)\frac{\partial\chi_1(y)}{\partial y_2}=0, \tag{13.6.31a}$$

①实际上 $C=\left(C_p\|b\|_{L_2}\right)^2$, 其中 C_p 是不等式 (2.4.7) 中的 Poincaré常数.

$$-D\Delta_y\chi_2(y)-b_2(y_1)\frac{\partial\chi_2(y)}{\partial y_2}=b_2(y_1) \tag{13.6.31b}$$

以及遍历性, 并且满足规范化条件, 即在单胞 $\mathcal{Y}$ 上 χ 积分为 0.

如果用 $\chi_1(y)$ 左乘第一个方程 (13.6.31a), 并在 $\mathbb{T}^d$ 使用分部积分, 则推得

$$\int_{\mathbb{T}^d}|\nabla_y\chi_1(y)|^2\mathrm{d}y=0.$$

由于 $\langle\chi(y)\rangle=0$, 其中 $\langle\cdot\rangle:=\int_{\mathbb{T}^d}\cdot\mathrm{d}y$, 因此, $\chi_1(y)=0$. 另一方面, 因为 (13.6.31b) 右边只依赖于 y_1, 所以假设解 $\chi_2(y)$ 独立于 y_2 是合理的. 假设找到一个具有上述形式的解, 则单胞问题解的唯一性暗含了其是唯一解. 方程 (13.6.31b) 变为

$$-D\frac{\mathrm{d}^2\chi_2(y_1)}{\mathrm{d}y_1^2}=b_2(y_1). \tag{13.6.32}$$

如果 ψ 是下面方程的一个周期解:

$$-\frac{\mathrm{d}^2\psi(y_1)}{\mathrm{d}y_1^2}=b_2(y_1), \tag{13.6.33}$$

则 ψ 独立于 D, 并且 $\chi_2=\psi/D$.

由 (13.6.24) 可知有效扩散系数 $\mathcal{K}$ 是下面的 2×2 阶矩阵:

$$\begin{aligned}\mathcal{K}&=\begin{pmatrix} D+\int_{\mathbb{T}^2}(b_1\chi_1)\,\mathrm{d}y & \frac{1}{2}\int_{\mathbb{T}^2}(b_2\chi_1+b_1\chi_2)\,\mathrm{d}y \\ \frac{1}{2}\int_{\mathbb{T}^2}(b_2\chi_1+b_1\chi_2)\,\mathrm{d}y & D+\int_{\mathbb{T}^2}(b_2\chi_2)\,\mathrm{d}y\end{pmatrix}\\ &=\begin{pmatrix} D & 0\\ 0 & \mathcal{K}_{22}\end{pmatrix},\end{aligned}$$

其中用到 $b_1=\chi_1=0$. 由 b_2,χ_2 仅依赖于 y_1 得到

$$\begin{aligned}\mathcal{K}_{22}:&=D+\int_0^1 b_2\chi_2\mathrm{d}y_1\\ &=D+\int_0^1-D\frac{\mathrm{d}^2\chi_2}{\mathrm{d}y_1^2}\chi_2\mathrm{d}y_1\\ &=D+D\int_0^1\left|\frac{\mathrm{d}\chi_2}{\mathrm{d}y_1}\right|^2\mathrm{d}y_1\\ &=D+\frac{1}{D}\int_0^1\left|\frac{\mathrm{d}\psi}{\mathrm{d}y_1}\right|^2\mathrm{d}y_1.\end{aligned}$$

值得注意的是, 由于 ψ 独立于 D, 则公式表明有效扩散系数随 D^{-1} 和初始分子扩散系数 D 一样趋于零, 这就表明在定理 13.19 中可以达到上界. 在这个尺度背后可以感知到对于很小的 D, 方程在 x_2 方向上可近似为一个运输方程. 运输的方向是慢慢调整变化的, 最终导致全局的扩散行为, 但是在长时间尺度上最显著的影响是运输效应, 这就增强了扩散性.

可以用作用在 b_2 上的积分算子来表示 ψ, 并且可说明

$$\mathcal{K}_{22} = D + \frac{1}{D}\|b_2\|^2_{H^{-1}_{\text{per}}(0,1)}, \tag{13.6.34}$$

见练习第 10 题.

例 13.20　考虑下面的情形:

$$b_2(y_1) = \sin(2\pi y_1). \tag{13.6.35}$$

使用式 (13.6.34) 并由练习第 10 题得

$$\mathcal{K}_{22} = D + \frac{1}{8\pi^2 D}. \tag{13.6.36}$$

图 13.2 中绘制了对数尺度上的有效扩散系数曲线, 该有效扩散系数可由 (13.6.36) 得到, 并且是分子扩散系数 D 的一个函数. 注意到当极限 $D \to 0$ 时, $\mathcal{K}$ 像 $1/D$ 一样发散. □

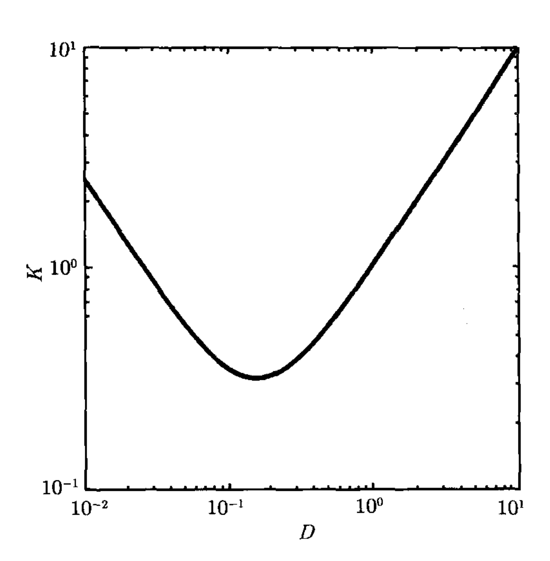


图 13.2　sine 剪切流中的等效扩散率对分子扩散率对数图

13.7 随机微分方程的联系

方程 (13.2.1) 是关于随机微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = b(x) + \sqrt{2D}\frac{\mathrm{d}W}{\mathrm{d}t} \tag{13.7.1}$$

的向后 Kolmogorov 方程, 其中 W 定义为 $\mathbb{R}^d$ 上的标准 Brownian 运动. 可以想到本章推导出的均匀化结论会对研究这个随机微分方程解的行为有所启示. 为此, 首先把由 (13.2.1) 推出 (13.2.4) 过程中用到的重新尺度标定的方法应用到随机微分方程 (13.7.1) 中, 即设

$$x \to \frac{x}{\varepsilon}, \quad t \to \frac{t}{\varepsilon^2},$$

则得

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{\varepsilon}b\left(\frac{x}{\varepsilon}\right) + \sqrt{2D}\frac{\mathrm{d}W}{\mathrm{d}t}. \tag{13.7.2}$$

(回想注 6.3 中关于重新标定时间尺度情况下白噪声行为.)

如果引入变量 $y = x/\varepsilon$, 则可将随机微分方程记为如下形式:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{1}{\varepsilon}b(y) + \sqrt{2D}\dfrac{\mathrm{d}W}{\mathrm{d}t}, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon^2}b(y) + \dfrac{1}{\varepsilon}\sqrt{2D}\dfrac{\mathrm{d}W}{\mathrm{d}t}, \end{cases}$$

其中当 y 在环面 $\mathbb{T}^d$ 上时, 将 x 视为 $\mathbb{R}^d$ 的一个元素. 这很类似于在第 11 章中分析过的形式 (11.2.1). 唯一不同的是出现在 x 和 y 中的噪声是相互关联的 (事实上, $U = V = W$). 在第 11 章中这改变了算子 $\mathcal{L}_1$, 因此, 不能将那章中推得的结论直接用到这里. 但是, 很容易地将其推广到研究相互关联的噪声上来, 见第 11 章中练习的第 1, 5 题. 注意到因为 ρ 是同一 Fokker-Planck 方程的稳态解, 所以中心条件 (13.2.8) 就是条件 (11.2.5).

本章的计算表明含有 (x, y) 变量的耦合随机微分方程的向后 Kolmogorov 方程可以用一个只含有唯一变量 x 的扩散方程来逼近. 实际上, 对于纯 Brownian 运动扩散方程就是向后 Kolmogorov 方程, 用随机微分方程解释得到下面的结论.

结论 13.21　假设中心条件 (13.2.8) 成立, 对于 $\varepsilon \ll 1$, 并且 $t = O(1)$, 随机微分方程 (13.7.2) 的解 x 可由下面方程的解 X 逼近:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \sqrt{(\mathcal{K} + \mathcal{K}^{\mathrm{T}})}\frac{\mathrm{d}W}{\mathrm{d}t},$$

其中由 (13.3.2) 得到矩阵 $\mathcal{K}$.

如果中心条件不满足, 则 (13.7.1) 中对尺度的重新标定可得到一个对流方程, 从而导致方程 (14.6.1), 该方程将在第 14 章中讨论.

13.8 讨论和参考

对于二阶抛物型偏微分方程均匀化问题及其与随机微分方程长时间渐近性的联系见文献 [33,ch.3], 其中还可以找到早期的参考文献, 也可见文献 [238]; 形如 (13.7.1) 的随机微分方程 (它的漂移项是一个周期标量函数的梯度) 描述了在周期势下的 Brownian 运动. 在许多应用中, 这是一个很重要的问题, 如在固态物理和生物学中, 见文献 [271,ch.11, 267] 及其参考文献. 在文献 [257] 中可以找到有关此类问题中多尺度方法应用方面的内容. 文献 [118, 238, 256, 323] 中也对梯度流的周期均匀化作了讨论. 在 [191] 中未采用任何均匀化理论, 推导出了一维势下移动的 Brownian 粒子的有效扩散系数的式 (13.6.13), 也可见文献 [138,sec. Ⅶ]. 在文献 [342] 中研究了一维情形双尺度周期位势下的 Brownian 运动. 多维问题的分析见文献 [258].

另外, 通常在流体扩散模型中出现具有无散漂移项的随机微分方程 (13.7.1). 周期的、不可压缩流的均匀化是湍流扩散理论的一部分, 见文献 [99, 200], 也可见文献 [100, 101, 221]. 因此, 出现了一个有趣的问题: 有效扩散系数依赖于分子扩散系数 D, 可以证明较小的 D 产生的有效扩散系数的渐近性敏感地依赖于流体速度场 $b(y)$ 的线流拓扑, 见文献 [20~22, 62, 63, 140, 294, 295]. 关于有效扩散率依赖于 D 而不是依赖于 Peclet 数 Pe 的有趣的试验结果见文献 [292, 293]; 重新标定尺度使得这些结论可用分子扩散系数来解释. 在文献 [223] 中研究了将可压缩流均匀化应用于分析大气输送现象.

即使当不满足中心条件 (13.2.8) 时, 也可能推导出一个均匀化方程. 此时, 有必要使用一个框架来一起移动, 一起移动的平均流为

$$\bar{b} = \int_{\mathbb{T}^d} b(y)\,\rho(y)\,\mathrm{d}y, \tag{13.8.1}$$

则对于重新标定尺度, 根据如下位移场可以推导出一个形如 (13.3.3) 的均匀化方程:

$$u^{\varepsilon}(x,t) = u\left(\frac{x}{\varepsilon} - \frac{\bar{b}t}{\varepsilon^2}, \frac{t}{\varepsilon^2}\right).$$

有效扩散系数由下面的公式给出:

$$\mathcal{K} = DI + 2D\int_{\mathbb{T}^d} \nabla_y \chi(y)^{\mathrm{T}} \rho(y)\,\mathrm{d}y + \int_{\mathbb{T}^d} \left(b(y) - \bar{b}\right) \otimes \chi(y)\rho(y)\,\mathrm{d}y. \tag{13.8.2}$$

单胞问题 (13.3.1) 变形为

$$-\mathcal{L}_0\chi = b - \bar{b}, \tag{13.8.3}$$

见第 14 章中练习的第 5 题.

对于周期性无散流, 依赖于很小的 D 的有效扩散系数的渐近性中, 平均流 $\bar{b}$ 有一个显著影响. 特别地, 对于含有 $D(D \ll 1)$ 的 $\mathcal{K}$ 的尺度, 其依赖于平均流是一个有理数还是一个无理数向量, 见文献 [35, 175, 201, 222, 295]. 关于非均匀性波长类似的有效扩散系数的间断依赖性可在文献 [126] 梯度流中观察到.

在 13.6.1 小节中证明了对于梯度流扩散性总是亏损的. 事实上, 可以得到更为准确的结论: 当 D 足够小时, 有效扩散系数较 D 是很小的, 即存在正常数 c_1 和 c_2, 使得

$$\langle \xi, \mathcal{K}\xi \rangle = c_1 \mathrm{e}^{-c_2/D}, \quad D \ll 1,$$

见 [54] 及其参考文献. 另一方面, 当把一个常外力添加到梯度漂移项上时, 并且当与分子扩散率进行比较时, 有效扩散率可以变得任意大, 见文献 [268, 282].

实际上, 沿着剪切方向有效扩散率与分子扩散率是成反比的 (式 (13.6.34)), 在文献 [313] 中没有使用任何均匀化理论对这个问题进行了讨论. 这个现象通常被称为 Taylor 扩散, 也可见文献 [11]. 通过直接求解, 对流扩散方程可以得到一个关于时变剪切流类似的结论, 见文献 [340].

为了推导出有效扩散率的表达式 (13.6.34)(来自练习第 10 题), 有必要使用 Fourier 级数进行计算. 当然, 必须证明可以对 Fourier 级数进行求导且 $\chi(y)$ 二阶导数的 Fourier 级数是有意义的. Fourier 级数的不同性质见文献 [132,ch.3].

已经证明对于梯度流, 其有效扩散张量是对称的. 但是, 对于一般向量场的有效扩散率不需要具有对称性. 尽管有效扩散率的反对称部分不会影响均匀化方程, 但是这具有重要的物理意义, 即会产生一个流量的组成部分, 其垂直于浓度的梯度, 见文献 [174]. 不管有效扩散率是否具有对称性, 或者是否不依赖于向量场 $b(y)$ 的对称性①, 在文献 [174, 253] 中, 针对无散流对这个问题进行了研究, 其参考文献中也研究了有效扩散率反对称部分对 Peclet 数的依赖性.

除了有效扩散率 (引起有效扩散张量 $\mathcal{K}$) 的 Eulerian 定义外, 也可以通过下面随机过程 $x(t)$ 方差的长时间平均定义一个 Lagrangian 有效扩散率:

$$D_{\mathrm{eff}}^{L} := \lim_{t\to\infty} \frac{\langle (x(t) - \langle x(t) \rangle) \otimes (x(t) - \langle x(t) \rangle) \rangle}{2t}. \tag{13.8.4}$$

注意: D_{eff}^{L} 是一个对称张量, 很容易证明 Lagrangian 有效扩散率 (13.8.4) 与 Eulerian 有效扩散率的对称部分是一致的.

多尺度方法也可应用于研究具有时变系数的抛物型偏微分方程的均匀化, 其中时变系数在 x 和 t 上也具有周期性, 见文献 [42, 118, 228, 257, 323].

① 例如, 因为梯度流的可逆性 (当然, 其就是一个对称特性), 所以有效扩散率在梯度流情形下是对称的.

在文献 [187] 中给出了关于对流扩散和运输偏微分方程的 Monte Carlo 方法. 用数值方法求解具有多尺度结构的对流方程的发展见文献 [1].

13.9 练　　习

1. 推导表达式 (13.4.1) 中第三项 $u_2(x, x/\varepsilon, t)$ 的公式.

2. 在一维情形下, 考虑均匀化问题

$$\frac{\partial u^\varepsilon}{\partial t} = -\frac{1}{\varepsilon}\nabla V\left(\frac{x}{\varepsilon}\right)\cdot\nabla u^\varepsilon + D\Delta u^\varepsilon,$$

其中势 (周期为 1) 为

$$V(y) = \begin{cases} y, & y \in \left[0, \dfrac{1}{2}\right], \\ 1-y, & y \in \left(\dfrac{1}{2}, 1\right], \end{cases}$$

计算有效扩散率 $\mathcal{K}$. 用 Laplace's 方法研究 $\mathcal{K}$ 依赖于小 D 的渐近性.

3. 根据第 2 题, 考虑势 $V(y) = \sin(2\pi y)$ 的情况. (提示: 使用 Bessel 函数.)

4. 在二维向量场 $b(y) = (b_1(y_1), b_2(y_1))$ 下计算有效扩散率 (13.3.3).

5. 考虑对流扩散反应方程的均匀化问题

$$\frac{\partial u^\varepsilon}{\partial t} = \frac{1}{\varepsilon}b\left(\frac{x}{\varepsilon}\right)\cdot\nabla u^\varepsilon + \Delta u^\varepsilon + \frac{1}{\varepsilon}c\left(\frac{x}{\varepsilon}\right)u^\varepsilon, \tag{13.9.1}$$

共中向量场 $b(y)$ 和尺度函数 $c(y)$ 是光滑的, 并且具有周期性. 使用多尺度方法使前面偏微分方程的均匀化. 特别地,

(a) 推导可解性条件;

(b) 得出可以推导出均匀化方程时 $b(y)$ 和 $c(y)$ 应该满足的条件;

(c) 推导均匀化方程、单胞问题以及均匀化系数的计算公式;

(d) 假设反应项是非线性的, 用下式代替方程 (13.9.1) 中的零阶项:

$$c\left(\frac{x}{\varepsilon}, u^\varepsilon\right),$$

其中对于所有 u, 函数 $c(y,u)$ 在 y 上的周期为 1. 此时, 可以将方程 (13.9.1) 均匀化吗?

6. 考虑随机微分方程的均匀化问题

$$\frac{\partial u^\varepsilon}{\partial t} = \left(b_1(x) + \frac{1}{\varepsilon}b_2\left(\frac{x}{\varepsilon}\right)\right)\cdot\nabla u^\varepsilon + \Delta u^\varepsilon, \tag{13.9.2}$$

其中向量场 $b_2(y)$ 是光滑的且具有周期性, $b_1(x)$ 也具有周期性. 使用多尺度方法使前面偏微分方程的均匀化, 特别地,

(a) 推导可解性条件;

(b) 得出可以推导出均匀化方程时 $b_2(y)$ 应该满足的条件;

(c) 证明均匀化方程为

$$\frac{\partial u}{\partial t} = b\cdot\nabla u + \mathcal{K} : \nabla\nabla u, \tag{13.9.3}$$

并且推导单胞问题以及均匀化系数 b 和 $\mathcal{K}$ 的计算公式.

7. 在如下情形:

$$b_1(x) = -\nabla V(x), \quad b_2(y) = -\nabla p(y),$$

考虑随机微分方程 (13.9.2) 的均匀化问题, 其中 $p(y)$ 具有周期性.

(a) 证明此时存在一个对称矩阵 $\hat{\mathcal{K}}$, 使得

$$\mathcal{K} = D\hat{\mathcal{K}}, \quad B = -\hat{\mathcal{K}}\nabla V;$$

(b) 设

$$\mathcal{L} := b \cdot \nabla + \mathcal{K} : \nabla\nabla u.$$

1. 推导 $\mathcal{L}^*$ 的计算公式, $\mathcal{L}^*$ 为 $\mathcal{L}$ 的 L^2 伴随;
2. 证明函数

$$\rho(y) := \frac{1}{Z}\mathrm{e}^{-V(y)/D}, \quad Z = \int_{\mathbb{T}^d} \mathrm{e}^{-V(y)/D}\mathrm{d}y$$

是均匀伴随函数

$$\mathcal{L}^*\rho = 0$$

的解.

8. 考虑下面随机微分方程的均匀化问题:

$$\frac{\partial u^\varepsilon}{\partial t} = b^\varepsilon \cdot \nabla u^\varepsilon + A^\varepsilon : \nabla_x\nabla_x u^\varepsilon,$$

其中 $A^\varepsilon = A(x/\varepsilon)$, $b^\varepsilon = b(x/\varepsilon)$, 向量场 $b(y)$ 以及矩阵 $A(y)$ 是光滑的且具有周期性, $A(y)$ 是正定的. 使用多尺度方法推导均匀化方程. 特别地,

(a) 推导可解性条件;

(b) 在 $b(y)$ 上得到确保均匀化方程存在的条件;

(c) 推导均匀化方程、单胞问题以及均匀化系数的计算公式;

(d) 证明均匀化矩阵是正定的.

9. 考虑下面随机微分方程的均匀化问题:

$$\frac{\partial u^\varepsilon}{\partial t} = \frac{1}{\varepsilon} b\left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2}\right) \cdot \nabla u^\varepsilon + D\Delta u^\varepsilon,$$

其中向量场 $b(y,\tau)$ 是光滑的、无散的, 并且在 y 和 τ 上周期为 1. 使用多尺度方法推导均匀化方程. 特别地,

(a) 推导可解性条件;

(b) 在 $b(y,\tau)$ 上得到确保均匀化方程存在的条件;

(c) 推导均匀化方程、单胞问题以及均匀化系数的计算公式;

(d) 证明均匀化矩阵是正定的.

10. 一个实值、周期为 1 的函数的 H_{per}^{-1} 范数可由下面的 Fourier 级数 (在 2.7 节中已经讨论过) 表示:

$$\|f\|_{H_{\mathrm{per}}^{-1}(0,1)}^2 = \frac{1}{2\pi^2}\sum_{k=1}^{\infty}\frac{|f_k|^2}{|k|^2}.$$

用此定义以及问题 (13.6.31b) 解 ψ 的 Fourier 级数表达式来建立式 (13.6.34) 关于 ψ 的 $\mathcal{K}_{22}$.

11. 当 $d=2$ 时, 在具有如下向量场的情形下考虑第 9 题:

$$b(y_1, y_2, \tau) = (0, b_2(y_1, \tau)).$$

推导有效扩散率 $\mathcal{K}$ 的计算公式, $\mathcal{K}$ 是如何依赖于 D 的?

12. 再次进行 13.6.2 小节中的计算以及二维向量场分别为

$$b(y_1, y_2) = (V, b(y_1))$$

和

$$b(y_1, y_2, \tau) = (V, b(y_1, \tau))$$

时的第 11 题, 其中 $V \in \mathbb{R}$. (提示: 使用方程 (13.8.2) 和 (13.8.3) 及其时变速度场的一般表达式.)

13. 设 $b(y)$ 是一个光滑的、实值的、周期为 1 的零值函数, 并且设 $\{b_k\}_{k=-\infty}^{+\infty}$ 是其 Fourier 系数, 证明 $b_0 = 0$ 以及 $b_{-k} = \bar{b}_k$.

第 14 章　线性双曲型和抛物型偏微分方程的平均化

14.1 引　　言

在本章中探讨了当中心条件不能满足时, 前面章节中讲到的线性双曲型 (或对流) 方程和抛物型 (或对流扩散) 方程解的长时间渐近行为. 采用的方法是文献中提到的均匀化方法. 但是, 按 1.3 节的分类方法实际上就是平均化方法, 仍然使用这个术语.

14.2 节和 14.3 节中建立了感兴趣的问题, 并陈述了近似结论; 14.4 节推导了当起始点是一个抛物型方程时的平均化方程; 14.5 节致力于完全运输方程情况下的平均化, 这个问题更精细 (对于首阶线性算子没有 Fredholm 性质), 这就是为什么专门投入一节研究它的原因; 14.6 节研究了常数微分方程和随机微分方程之间的关系; 14.7 节给出了一些文献综述.

14.2 完 全 方 程

在一个稳态周期速度场 b 下, 研究如下线性抛物型方程:

$$\frac{\partial u}{\partial t} = b \cdot \nabla u + D\Delta u, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}^+, \tag{14.2.1a}$$

$$u = u_{\text{in}}, \quad (x,t) \in \mathbb{R}^d \times \{0\} \tag{14.2.1b}$$

的解的长时间渐近行为, 其中这个抛物型方程描述了对流扩散过程, 并且在 $D=0$ 情况下, 它退化为一个线性运输方程. 如第 13 章中研究的情况, 其中:

$$u_{\text{in}}(x) = g(\varepsilon x),$$

在空间和时间上采用新的尺度重新推导方程以了解方程 (14.2.1) 在长度和时间尺度上解的性能, 这些时间尺度与速度场 $b(x)$ 的尺度相比要长得多. 在这个背景下, 这个问题的小参数是速度场中特征长度 (时间) 的尺度 (它的周期) 和这个问题中的最大长度 (时间) 的尺度之间的比值, 需要对这个问题寻求一个平均化的描述. 对照前面章节中对流扩散方程的分析, 在相同的方式下重新建立时间和空间尺度, 即

$$x \to \varepsilon^{-1}x, \quad t \to \varepsilon^{-1}t. \tag{14.2.2}$$

在抛物型 $D>0$ 的情况下, 因为没有假设中心条件 (13.2.8) 成立, 从而平流作用达到平均. 由于运输偏微分方程 (14.2.1a) 在空间和时间都是一阶的, 则当 $D=0$ 时, 运输方程 (14.2.2) 也是很自然的.

要研讨的初始值问题变为

$$\frac{\partial u^\varepsilon}{\partial t} = b^\varepsilon \cdot \nabla u^\varepsilon + \varepsilon D \Delta u^\varepsilon, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}^+, \tag{14.2.3a}$$

$$u^\varepsilon = f, \quad x \in \mathbb{R}^d \times \{0\}. \tag{14.2.3b}$$

如前面章节中, $b^\varepsilon(x) = b(x/\varepsilon)$.

如前面章节中, 定义算子

$$\mathcal{L}_0 = b\,(y) \cdot \nabla_y + D\Delta_y \tag{14.2.4}$$

具有周期边界条件. y 中的常数在这个算子的零空间上, 此外, 对于 $D>0$ 这个零空间是一维的且仅由常数构成. $\mathcal{L}_0$ 的 L^2 伴随是 $\mathcal{L}^*$, 也具有周期边界条件. 回顾第 13 章, 对于 $D>0$, **不变分布**$\rho(y)$ 是伴随方程的唯一稳态解,

$$\mathcal{L}_0^*\rho = 0, \quad \int_{\mathbb{T}^d} \rho(y)\mathrm{d}y = 1 \tag{14.2.5}$$

且有周期边界条件. 对于 $D>0$, 算子 $\mathcal{L}_0$ 和 $\mathcal{L}^*$ 都满足一个 Fredholm 性质.

14.3 简化方程

在本节和下一节中简单地设算子 $\mathcal{L}_0$ 有一个由常数构成的一维零空间, 并且对于它的伴随 $\mathcal{L}_0^*$ 同样成立, 这个零空间是由 ρ 张成的. 对于 $D>0$, 遵循 Fredholm 性质; 对于 $D=0$, 要求常微分方程的一些遍历性形式, 对于这个方程来说, $\mathcal{L}_0$ 是生成子; 将在 14.5 节和 14.6 节中讨论遍历性问题.

在上述关于 $\mathcal{L}_0$ 假设的情况下有如下结论:

结论 14.1 设 b 是一个光滑周期向量场, (14.2.4) 中定义的算子 $\mathcal{L}_0$ 满足

$$\mathcal{N}(\mathcal{L}_0) = \mathrm{span}(1), \quad \mathcal{N}(\mathcal{L}_0^*) = \mathrm{span}(\rho),$$

则对于 $\varepsilon \ll 1$ 且 t 是 $O(1)$ 的, (13.2.4) 的解 $u^\varepsilon(x,t)$ 近似于 $u(x,t)$, 平均化方程

$$\frac{\partial u}{\partial t} - \bar{b} \cdot \nabla_x u = 0, \quad \bar{b} := \int_{\mathbb{T}^d} \rho(y) b(y) \mathrm{d}y$$

的解与 u^ε 有一样的初始条件.

导致这个近似结论的计算是通过采用新的尺度对抛物型方程 (14.2.3a) 进行推导作为出发点, 并通过平均化的方式重新获得一个运输方程. 如果认为由于乘以 ε, (14.2.3a) 中的扩散项可能很简单地从平均化计算中消失, 这个观点是错误的, 因为扩散系数扮演着一个很必要的角色. 一般来说, 通过 ρ①, 稳态分布的形式关键取决于 $D>0$.

注意到中心条件 (13.2.8) 简单地说明了 $\bar{b}=0$. 这就是第 13 章用到的时间和空间尺度与这里采用的不同的原因, 尤其是这里采用了更长的时间尺度, 以看到其不可忽视的作用.

14.4 推 导

使用前两章中介绍的多尺度方法来介绍辅助变量 $y=x/\varepsilon$. 设 $\phi=\phi(x,x/\varepsilon)$ 是标量值. 链式法则给定

$$\nabla\phi=\nabla_x\phi+\frac{1}{\varepsilon}\nabla_y\phi \quad 和 \quad \Delta\phi=\Delta_x\phi+\frac{2}{\varepsilon}\nabla_x\cdot\nabla_y\phi+\frac{1}{\varepsilon^2}\Delta_y\phi.$$

方程 (14.2.3) 右边出现的偏微分算子有如下形式:

$$\mathcal{L}=\frac{1}{\varepsilon}\mathcal{L}_0+\mathcal{L}_1+\varepsilon\mathcal{L}_2,$$

其中

$$\begin{aligned}
\mathcal{L}_0 &= b(y)\cdot\nabla_y+D\Delta_y,\\
\mathcal{L}_1 &= b(y)\cdot\nabla_x+2D\nabla_x\cdot\nabla_y,\\
\mathcal{L}_2 &= D\Delta_x.
\end{aligned}$$

根据 x 和 y, 方程 (14.2.3) 变为

$$\frac{\partial u^\varepsilon}{\partial t}=\left(\frac{1}{\varepsilon}\mathcal{L}_0+\mathcal{L}_1+\varepsilon\mathcal{L}_2\right)u^\varepsilon.$$

为了寻找双尺度展开式

$$u^\varepsilon(x,t)=u_0\left(x,\frac{x}{\varepsilon},t\right)+\varepsilon u_1\left(x,\frac{x}{\varepsilon},t\right)+\cdots \tag{14.4.1}$$

的形式解, 假定展开式 $u_j=(x,y,t)(j=0,1,\cdots)$ 中所有的项在 y 中都是 1 周期的, 并将 x 和 $y:=x/\varepsilon$ 看成独立变量②. 把 (14.4.1) 代入方程 (14.2.3a), 使用假设 x 和

① 无散度流的情况是一个例外, 对于所有的 $D>0$, 不变测度 ρ 是单位环面上的 Lebesgue 测度, 见定理 13.15.

② 正如前面两章在椭圆型和抛物型均匀化过程中, 这就是利用尺度分离的地方: 将 x 和 y 看成是独立变量. 当 $\varepsilon\to 0$ 时对这个假设的验证是平均化严格理论的重要问题, 见第 21 章.

y 相互独立, 根据 ε 的幂次合并同类项, 则可得如下方程组:

$$O\left(\frac{1}{\varepsilon}\right):-\mathcal{L}_0 u_0=0; \tag{14.4.2a}$$

$$O(1):-\mathcal{L}_0 u_1=\mathcal{L}_1 u_0-\frac{\partial u_0}{\partial t}, \tag{14.4.2b}$$

其中 $u_j(x,y)$ 在 y 上是 1 周期的.

接下来可以完成平均化过程. 从 (14.4.2) 的第一个方程和对 $\mathcal{L}_0$ 的假设, 可以推导出展开式第一项独立于振荡, 这一个振荡可以通过辅助变量 y 表示, 则 u_0 可表示为

$$u_0=u(x,t).$$

运用这个可以计算

$$\mathcal{L}_1 u_0=\frac{\partial u(x,t)}{\partial t}-b(y)\cdot\nabla_x u(x,t).$$

由于 ρ 在 $\mathcal{L}_0^*$ 的零空间上, (14.4.2) 的第二个方程暗含了

$$0=\frac{\partial u(x,t)}{\partial t}-\left(\int_{\mathbb{T}^d}\rho(y)b(y)\mathrm{d}y\right)\cdot\nabla_x u(x,t), \tag{14.4.3}$$

这样就得到了想要的平均化方程

$$\frac{\partial u(x,t)}{\partial t}-\bar{b}\cdot\nabla_x u(x,t)=0,\quad \bar{b}:=\int_{\mathbb{T}^d}\rho(y)b(y)\mathrm{d}y,$$

并且具有与 u^ε 相同的初始条件.

14.5 运输方程：$D=0$

一般地, 平均化运输方程通过不变分布 ρ(在这个分布上, b 是取平均的) 主要取决于扩散系数 D. 当 $D>0$ 时, 稳态分布 ρ 的存在性和唯一性是自动的, 但是当 $D=0$ 时, 则要求一些遍历性形式, 这将取决于 b 的性质. 因此, 会问到一个问题, 即什么时候和在什么程度下, 前面的平均化计算可以拓展到 $D=0$ 的情况; 这些计算所依赖的 $\mathcal{L}_0$ 和 $\mathcal{L}_0^*$ 的零空间是否为一维的, 当 $D>0$ 时, 有时可以通过 Fredholm 性质来加以保证. 下面来讨论 $D=0$ 时的这些类似结论.

设

$$\mathcal{L}_0=b(y)\cdot\nabla_y \tag{14.5.1}$$

有定义域 $C^1_{\mathrm{per}}(\mathbb{T}^d)$, 如 (4.3.8) 可以将这个算子拓展到 $\mathcal{D}(\mathcal{L}_0)\subset L^\infty_{\mathrm{per}}(\mathbb{T}^d)$. 设在 $\mathcal{L}_0$ 的零空间 $\mathcal{N}$ 中没有非平凡函数. 即

$$\mathcal{N}(\mathcal{L}_0)=\{y\text{上的常数}\}. \tag{14.5.2}$$

$\mathcal{D}(\mathcal{L}_0)$ 是 $\mathcal{L}_0$ 的定义域空间. 由第 4 章中可知, 对具有向量场 b 的常微分方程给定一个遍历性假设是非常必要的, 见定理 4.13(iii). 关于这个, $\mathcal{L}_0^*$ 是非空的, 其定义域被认为是 $C^1_{\text{per}}(\mathbb{T}^d)$, 这个观点暗含了一个不变测度的存在, 并且这个测度关于 Lebesgue 测度是绝对连续的, 见定理 4.12(iii), 从而关于绝对连续不变测度 μ 的遍历性, 提供了一个有用的工具去实现本章中 $D=0$ 的情况下进行的形式扰动展开. 特别地, 在遍历性情况下, (14.4.2a) 暗含了 u_0 是独立于 y 的, 也暗含了 (14.4.2b) 的解 u_1 存在的必要条件是方程 (14.4.3).

注意到如果 b 是无散的 (速度场是不可压缩的), 则 $\mathcal{L}$ 是反对称矩阵 (引理 13.14), 所以可以从 (14.5.2) 推导出

$$\mathcal{N}(\mathcal{L}_0^*) = \{y\text{上的常数}\}. \tag{14.5.3}$$

然而, 在一般的遍历性情况下, ρ 并不是一个常数函数.

14.5.1　一维例子

在一维中考虑运输方程 (14.2.3a)

$$\frac{\partial u^\varepsilon}{\partial t} - b^\varepsilon \frac{\partial u^\varepsilon}{\partial x} = 0, \quad (x,t) \in \mathbb{R} \times \mathbb{R}^+, \tag{14.5.4a}$$

$$u = g, \quad (x,y) \in \mathbb{R} \times \{0\}, \tag{14.5.4b}$$

其中 $g = g(x)$ 独立于振荡[①]. 设 $b(y)$ 是严格正的、光滑的、1 周期函数. 稳态 Liouville 方程

$$\mathcal{L}_0^* \rho = 0, \quad \rho > 0, 1\text{周期}, \tag{14.5.5}$$

具有标准规范化条件

$$\int_0^1 \rho(y)\mathrm{d}y = 1,$$

并且有唯一规范的概率密度解:

$$\rho(y) = \frac{C}{b(y)}, \quad C = \left\langle b(y)^{-1} \right\rangle^{-1}. \tag{14.5.6}$$

如第 12 章中, 在这里使用记号 $\langle\cdot\rangle$ 表示 [0,1] 上的平均化. b 的正性是结论存在且唯一的关键, 也是基本流遍历性的关键. 这些问题的讨论见例 4.14.

正如在 (14.5.4b) 相同的初始条件有

$$\bar{b} = \left\langle b(y)^{-1} \right\rangle^{-1},$$

① 这并不是必要条件, 见第 21 章练习第 3 题.

可得平均化方程

$$\frac{\partial u}{\partial t}-\bar{b}\frac{\partial u}{\partial x}=0. \tag{14.5.7}$$

注意到在下一节中与遍历性无散度相比的情况，它是出现在平均化方程 (14.5.7) 速度场的调和平均值中，而不是标准平均值. (调和平均值也同样出现在一维椭圆型情况下，见 12.6.1 小节.)

14.5.2 无散度速度场

如果 b 是无散的 (速度场是不可压缩的)，则由 (14.5.1) 给出 $\mathcal{L}$ 是反对称阵 (引理 13.14)，因此，可推导出如果 (14.5.2) 成立，则

$$\mathcal{N}(\mathcal{L}_0^*)=\{y\text{上的常数}\} \tag{14.5.8}$$

(见例 4.15). 但是，甚至是对无散场，导致 (14.5.2) 的遍历性假设也不能满足. 考虑方程 (14.4.2a) 的如下形式：

$$\mathcal{L}_0 u:=b(y)\cdot\nabla_y u=0. \tag{14.5.9}$$

方程 (14.5.9) 具有周期边界条件. 尽管 $u\equiv 1$ 是这个方程的解，但很少见这个解是唯一的. 一般来说，算子 $\mathcal{L}_0$ 的零空间包含 y 的非平凡函数. 举一个例子，考虑一个光滑、1 周期、无散场

$$b(y)=(\sin(2\pi y_2),\sin(2\pi y_1)).$$

很容易证明函数

$$u(y)=\cos(2\pi y_1)-\cos(2\pi y_2)$$

可解方程 (14.5.9). 因此，$\mathcal{L}_0$ 的零空间依赖于速度场 $b(y)$，一般说它不包括 y 中的常数，这也暗含了不能用多尺度方法实现平均化过程.

很自然地会问是否存在一个方法可以判别 $\mathbb{T}^d$ 上给出的速度场是否具有遍历性，这在二维空间上也是可能的，沿着这些思路可得如下结论：

定理 14.2 设 $b(y):\mathbb{T}^2\to\mathbb{R}^2$ 是一个光滑、无散的速度场，满足

$$b_1(y)\neq 0,\quad \forall y\in\mathbb{T}^2,$$

使得它没有滞留点. 设 $\bar{b}_i(i=1,2)$ 表示 $\mathbb{T}^2$ 上速度场的第 i 分量，并定义转动数为

$$\gamma=\frac{\bar{b}_1}{\bar{b}_2},$$

则变量 $y\mapsto z$ 存在一个平稳变化. 在这个变化下，常微分方程

$$\frac{\mathrm{d}y_1}{\mathrm{d}t}=b_1(y),\quad \frac{\mathrm{d}y_2}{\mathrm{d}t}=b_2(y) \tag{14.5.10}$$

转化为

$$\frac{\mathrm{d}z_1}{\mathrm{d}t} = g(z), \quad \frac{\mathrm{d}z_2}{\mathrm{d}t} = \gamma g(z), \tag{14.5.11}$$

其中 $g(z)$ 是非零且光滑的标量函数. 此外, 如果设 γ 是无理数, 则算子 $\mathcal{L}_0$ 的零空间在 $D(\mathcal{L}_0)$ 上是一维的.

证明 定理的第一部分可通过构造显式变化, 把 (14.5.10) 映射到 (14.5.11)①,

$$z_1 = \frac{1}{\bar{b}_2}\int_0^{y_1} b_2(\xi, 0)\mathrm{d}\xi, \quad z_2 = \frac{1}{\bar{b}_1}\int_0^{y_2} b_1(y_1, \xi)\mathrm{d}\xi.$$

定理的第二部分可以通过 Fourier 分析证明, 见练习第 7 题. □

因此, 在定理 14.2 的条件下, 定理 4.13 成立, 并且也可以得到本章中的形式扰动展开式.

14.6 常微分方程和随机微分方程的联系

首先考虑 $D = 0$ 的情况. 回顾 (14.2.3) 的解是由

$$u(x, t) = g(\varphi^t(x))$$

给定的, 其中 $\varphi^t(x)$ 是如下常微分方程的解:

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi^t(x) = b\left(\frac{\varphi^t(x)}{\varepsilon}\right),$$

$$\varphi^t(x) = x.$$

结论 14.1 说明了当遍历性假设成立时, $\mathcal{L}_0$ 有一个一维零空间, 这个方程也可以由如下式子近似逼近:

$$\overline{\varphi}^t(x) = \bar{b}t + x,$$

$\overline{\varphi}^t(x)$ 是如下方程的解:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\varphi}^t(x) = \bar{b},$$

$$\overline{\varphi}^t(x) = x,$$

其中

$$\bar{b} = \int_{\mathbb{T}^d} \rho(y) b(y) \mathrm{d}y.$$

由 (14.5.6) 可知在一维 ($d = 1$) 中有

$$\bar{b} = \left\langle b(y)^{-1} \right\rangle.$$

① 在额外假设条件 $b_2 \neq 0$ 下, 这在后期是可以消除的.

另一个得到这个结论的方法如下: 设 $x=\varphi^t(x_0)$ 和 $y=x/\varepsilon$, 则

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= b(y),\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{1}{\varepsilon}b(y).\end{aligned}$$

在遍历性假设下, 快速过程 y 在圆环 $\mathbb{T}^d$ 上有一个不变测度 ρ, 从而平均化结论 10.1 给出了 x 的近似可由如下方程给出:

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\overline{b},$$

例 14.3 在一维情况下, 运用特征法推导平均化方程也是可能的. 为了得到这个结论, 在一维中和前面相同的假设下, 考虑方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=b\left(\frac{x}{\varepsilon}\right).$$

如果令 $y=x/\varepsilon$, 则可直接得到

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{1}{\varepsilon}b(y),$$

从而, 如果通过

$$T=\int_0^1\frac{1}{b(z)}\mathrm{d}z=\frac{1}{\overline{b}}$$

定义 T, 则

$$y(n\varepsilon T)=\frac{x(0)}{\varepsilon}+n.$$

因此,

$$x(n\varepsilon T)=x(0)+n\varepsilon.$$

从 $x(t)$ 收敛于 $X(t)$ 的连续性可以推出上述式子, 其中

$$X(t)=x(0)+\frac{t}{T}.$$

这个极限函数 $X(t)$ 满足均匀化方程

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\frac{1}{T}=\overline{b}.$$ □

如果 $D>0$, 则对于随机微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=b\left(\frac{x}{\varepsilon}\right)+\sqrt{2\varepsilon D}\frac{\mathrm{d}W}{\mathrm{d}t},$$

方程 (14.2.3) 是它的向后 Kolmogorv 方程. 另一个解释平均化结论的方法如下: 设 $y=x/\varepsilon$, 可得

$$\frac{\mathrm{d}x}{\mathrm{d}t}=b(y)+\sqrt{2\varepsilon D}\frac{\mathrm{d}W}{\mathrm{d}t},$$
$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{1}{\varepsilon}b(y)+\sqrt{\frac{2D}{\varepsilon}}\frac{\mathrm{d}W}{\mathrm{d}t}. \tag{14.6.1}$$

在遍历性假设下, 快速过程 y 在圆环 $\mathbb{T}^d$ 上有一个不变测度 ρ, 从而平均化结论 10.1 的推广给出了 x 可以由如下常微分方程近似:

$$\frac{\mathrm{d}x}{\mathrm{d}t}=\bar{b}.$$

14.7　讨论和参考

这里运用的扰动展开式是类似于第 9 章和第 10 章中 Markov 链中常微分方程和随机微分方程的平均化方法. 线性运输方程的平均化问题被很多学者研究, 见文献 [51, 80, 147, 312,]; 随机微分方程平均化问题的详细研究见文献 [111].

当 $D=0$ 且仅当速度场具有遍历性时, 多尺度方法可以获得平均化的线性运输方程 (14.2.3a); 当速度场不具有遍历性时, 多尺度方法失效, 因为在这种种情况下, 没有可以进行平均化的可解条件. 实际上, 当速度场不具有遍历性时, $\varepsilon\to 0$ 极限变得更复杂, 并且极限过程也不能通过一个简单的随机方程表达. 为了研究一般的速度场问题, 无需具有遍历性, 运用双尺度收敛方法也是可能的, 这将在第 21 章中讲到.

定理 14.2 的证明见文献 [312], 其结论的证明不仅是针对无散流 (其不变密度简单设为 1), 也是针对具有光滑不变密度的二维流. 这不是一个可以证明的结论, 而是假设 $b_1\neq 0$ 所导致的一个特别简单的证明. 在假设无滞留点情况下类似的证明可见文献 [290]. 在 $\mathbb{T}^2$ 上具有不变测度, 而不是 Lebesgue 测度的速度场, 其类似的定理也是成立的, 见文献 [312].

14.5.1 小节中例题的研究见文献 [81,312]. 对流扩散及运输偏微分方程的 Monte Carlo 方法的提出见文献 [187].

14.8　练　　习

1. 如果允许 b 更换符号, 则 14.5.1 小节中是如何研究常微分方程的动态性的?

2. 在一维情况下, 假设 a(或 b) 是光滑 1 周期的以及 $\inf_x a>0$(或 $\inf_x b>0$) 情况下考虑方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=a\left(\frac{x}{\varepsilon}\right)b\left(\frac{t}{\varepsilon^{\alpha}}\right),$$

找出平均化方程.

3. 研究 (14.2.3) 中的平均化问题, 说明它具有一个光滑周期 (剪切) 速度场 $b:\mathbb{T}^2\mapsto\mathbb{R}^2$, 并且形式为

$$b(y)=(0,b_2(y_1))^{\mathrm{T}}.$$

4. 研究 (14.2.3) 中的平均化问题, 说明它具有一个速度场 $b:\mathbb{T}^2\mapsto\mathbb{R}^2$, 并且形式为

$$b(y)=\hat{b}(y)(0,\gamma)^{\mathrm{T}},$$

其中 $\hat{b}(y)$ 是一个光滑、1 周期标量函数且 $\gamma\in\mathbb{R}$.

5. 考虑方程 (13.2.4), 其中心条件 (13.2.8) 不成立, 说明对于重新标定尺度

$$u^{\varepsilon}(x,t)=u\left(\frac{x}{\varepsilon}-\frac{\bar{b}t}{\varepsilon^2},\frac{t}{\varepsilon^2}\right),$$

可以推导出一个 (13.3.3) 形式的均匀化方程, 其中 u 是 (14.2.1) 的解, $\bar{b}$ 是由结论 14.1 给出的. 说明单胞问题变成

$$-\mathcal{L}_0\chi=b-\bar{b}, \tag{14.8.1}$$

并且有效扩散率是由如下方程给出:

$$\mathcal{K}=DI+2D\int_{\mathbb{T}^d}\nabla_y\chi(y)^{\mathrm{T}}\rho(y)\mathrm{d}y+\int_{\mathbb{T}^d}\left(\left(b(y)-\bar{b}\right)\otimes\chi(y)\rho(y)\right)\mathrm{d}y. \tag{14.8.2}$$

6. 研究常微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=-\nabla V\left(\frac{x}{\varepsilon}\right)+F$$

的均匀化问题, 其中 $V(y)$ 是光滑的、1 周期函数, F 是一个常向量.

7. 完成定理 14.2 的详细证明.



第三部分　理　　论

第 15 章　常微分方程的不变流形：收敛性定理

15.1　引　　言

在本章中给出了一种严格理论, 它证实了第 8 章中不变流形的扰动展开式. 15.2 节讲述了一个近似定理; 在 15.3 节中运用常微分方程理论的基本方法, 证明了这样一个直接估计, 它在任意有限时间区域上是有效的; 在 15.4 节中, 提到了在任意时间间隔上建立相应结论更为熟练的技巧.

15.2　定　　理

对于 $\varepsilon \ll 1$ 和 $x \in \mathbb{R}^l, y \in \mathbb{R}^{d-l}$, 研究方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x,y), & (15.2.1\mathrm{a}) \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{1}{\varepsilon} g(x,y). & (15.2.1\mathrm{b}) \end{cases}$$

假定当 x 固定时, y 的动态性有唯一指数形式的吸引性固定点 (在 x 上一致). 特别地, 假设存在 $\eta : \mathbb{R}^l \to \mathbb{R}^{d-l}$ 和 $\alpha > 0$, 使得对于所有的 $x \in \mathbb{R}^l$ 和所有的 $y_1, y_2 \in \mathbb{R}^{d-l}$,

$$g(x, \eta(x)) = 0, \tag{15.2.2a}$$

$$\langle g(x,y_1) - g(x,y_2), y_1 - y_2 \rangle \leqslant -\alpha |y_1 - y_2|^2. \tag{15.2.2b}$$

将参考 (15.2.2b) 作为压缩性条件.

当 x 固定为 ξ 时的动态性满足

$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi_\xi^t(y) = g\left(\xi, \varphi_\xi^t(y)\right), \quad \varphi_\xi^0(y) = y. \tag{15.2.3}$$

设向量场 $(f^{\mathrm{T}}, g^{\mathrm{T}})^{\mathrm{T}}$ 是 $\mathbb{R}^d$ 上的 Lipschitz, 则可使得 (15.2.1) 和 (15.2.3) 都有全局定义解.

对 η 和 g 的假设暗含了 $\varphi_\xi^t(y)$ 的如下指数收敛于它的整体吸引固定点 $\eta(\xi)$.

引理 15.1　在 (15.2.2) 假设下可知, 对于所有的 $y \in \mathbb{R}^{d-l}$ 有

$$|\varphi_\xi^t(y) - \eta(\xi)| \leqslant \mathrm{e}^{-\alpha t} |y - \eta(\xi)|.$$

证明　由 $\eta(\xi)$ 具有时变性可知

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_\xi^t(y) = g(\xi, \varphi_\xi^t(y)),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta(\xi) = g(\xi, \eta(\xi)) = 0.$$

因此,

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\varphi_\xi^t(y) - \eta(\xi)|^2 &= \langle g(\xi, \varphi_\xi^t(y)) - g(\xi, \eta(\xi)), \varphi_\xi^t(y) - \eta(\xi)\rangle \\ &\leqslant -\alpha|\varphi_\xi^t(y) - \eta(\xi)|^2.\end{aligned}$$

其结论遵从 Gronwall 引理 4.4 的微分形式.

实质上, 希望证明这样一个结论, 当 x 不再固定为 ξ, 而是在自身的时间尺度上演化时, 用 $y(t)$ 代替 $\varphi_\xi^t(y)$, 这相对于 y 的同一演化是很缓慢的. 作如下假设, 这些假设能简化分析, 并能使证明基本结论的方法更加清晰. 然而, 在各种不同方式下它们都可以被减弱. 这些假设就是存在一个常数 $C > 0$, 使得①

$$|f(x,y)| \leqslant C, \quad \forall (x,y) \in \mathbb{R}^d; \tag{15.2.4a}$$

$$|\nabla_x f(x,y)| \leqslant C, \quad \forall (x,y) \in \mathbb{R}^d, \tag{15.2.4b}$$

$$|\nabla_y f(x,y)| \leqslant C, \quad \forall (x,y) \in \mathbb{R}^d, \tag{15.2.4c}$$

$$|\eta(x)| \leqslant C, \quad \forall x \in \mathbb{R}^l, \tag{15.2.4d}$$

$$|\nabla \eta(x)| \leqslant C, \quad \forall x \in \mathbb{R}^l. \tag{15.2.4e}$$

运用这些假设可以证明 x 是

$$\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, \eta(X)), \quad X(0) = x(0) \tag{15.2.5}$$

的一个近似于 X 的解.

定理 15.2　假定假设 (15.2.2) 和 (15.2.4) 成立, 则存在常数 K, $c > 0$, 使得 $x(t)$ 是 (15.2.1) 的解, $X(t)$ 是 (15.2.5) 的解, 并满足

$$|x(t) - X(t)|^2 \leqslant c\mathrm{e}^{Kt}\left(\varepsilon|y(0) - \eta(x(0))|^2 + \varepsilon^2\right).$$

注意到误差是 $\sqrt{\varepsilon}$, 它是 $O(1)$. 如果初始偏差 $y(0) - \eta(x(0))$ 是 $\sqrt{\varepsilon}$, 那么这也可以减小到 ε. 此外, 注意到对于给出的 ε, 定理 15.2 仅在 $\ln \varepsilon^{-1}$ 阶次的区间上有用. 更有效的结论也是可能的, 并将在 15.4 节中加以探讨.

① 这些条目中第 1 项和第 4 项的范数都是标准向量范数; 在第 2, 3, 5 项中它们都是矩阵 (算子) 范数. 所有的都是 Euclidean. 事实上, η 流假设性能是源于对 g 的适当假设, 这是由于 x 中压缩假设 (15.2.2b) 的一致性.

15.3 证　明

由 $y(t)=\eta(x(t))+z(t)$ 定义 $z(t)$, 则运用假设 (15.2.2a) 可得

$$\begin{aligned}\frac{\mathrm{d}z}{\mathrm{d}t}&=\frac{\mathrm{d}y}{\mathrm{d}t}-\nabla\eta(x)\frac{\mathrm{d}x}{\mathrm{d}t}\\&=\frac{1}{\varepsilon}g(x,\eta(x)+z)-\nabla\eta(x)f(x,\eta(x)+z)\\&=\frac{1}{\varepsilon}(g(x,\eta(x)+z)-g(x,\eta(x)))-\nabla\eta(x)f(x,\eta(x)+z),\end{aligned}$$

接着假设 (15.2.2b) 暗含了

$$\langle g(x,\eta(x)+z)-g(x,\eta(x)),z\rangle\leqslant-\alpha|z|^2.$$

此外, 假设 (15.2.4) 暗含了

$$\langle z,\nabla\eta(x)f(x,\eta(x)+z)\rangle\leqslant C^2|z|.$$

因此, 运用 Cauchy-Schwarz 不等式中 $\delta^2=\varepsilon/\alpha$ 可得

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|z|^2&=\left\langle z,\frac{\mathrm{d}z}{\mathrm{d}t}\right\rangle\\&=\frac{1}{\varepsilon}\langle z,g(x,\eta(x)+z)-g(x,\eta(x))\rangle-\langle z,\nabla\eta(x)f(x,\eta(x)+z)\rangle\\&\leqslant-\frac{\alpha}{\varepsilon}|z|^2+C^2|z|\\&\leqslant-\frac{\alpha}{\varepsilon}|z|^2+\frac{\delta^2}{2}C^4+\frac{1}{2}\frac{|z|^2}{\delta^2}\\&\leqslant-\frac{\alpha}{2\varepsilon}|z|^2+\frac{\varepsilon}{2\alpha}C^4.\end{aligned}$$

由 Gronwall 引理可得

$$|z(t)|^2\leqslant\mathrm{e}^{-\frac{\alpha}{\varepsilon}t}|z(0)|^2+\left(1-\mathrm{e}^{-\frac{\alpha}{\varepsilon}t}\right)\frac{\varepsilon^2C^4}{\alpha^2}. \tag{15.3.1}$$

接着

$$\frac{\mathrm{d}X}{\mathrm{d}t}=f(X,\eta(X))$$

和

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x,\eta(x)+z)$$

相减可得

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(x-X)&=f(x,\eta(x)+z)-f(X,\eta(X))\\&=f(x,\eta(x)+z)-f(x,\eta(X))+f(x,\eta(X))-f(X,\eta(X)).\end{aligned}$$

接下来运用 Cauchy-Schwarz 不等式以及 $f(\cdot,\cdot)$ 和 $\eta(\cdot)$ 的 Lipschitz 连续估计,

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|x-X|^2 &= \left\langle x-X, \frac{\mathrm{d}}{\mathrm{d}t}(x-X)\right\rangle \\ &= \langle x-X, f(x,\eta(x)+z)-f(x,\eta(X))\rangle \\ &\quad +\langle x-X, f(x,\eta(X))-f(X,\eta(X))\rangle \\ &\leqslant C|\eta(x)-\eta(X)+z||x-X|+C|x-X|^2 \\ &\leqslant (C^2+C)|x-X|^2+C|z||x-X|,\end{aligned}$$

可得

$$\frac{\mathrm{d}}{\mathrm{d}t}|x-X|^2 \leqslant (3C^2+2C)|x-X|^2+|z|^2.$$

设 $K=3C^2+2C$, 并运用 $|z(t)|^2$ 的边界 (15.3.1), Gronwall 引理以及 $x(0)=X(0)$ 的事实可以得到想要的结论. □

15.4 讨论和参考

定理 15.2 显示了完全方程 (15.2.1) 的 $x(t)$ 仍然接近于在时间尺度上简化方程 (15.2.5) 的解 $X(t)$, 这个时间尺度的阶数为 $\ln(\varepsilon^{-1})$. 在更长时间尺度上的单个解是发散的, 这是因为在任何动力系统中都可以表现出轨迹的指数分离. 然而注意到估计 (15.3.1) 说明了

$$\limsup_{t\to\infty}|y(t)-\eta(x(t))| \leqslant \frac{C^4}{\alpha^2}\varepsilon^2,$$

它表明在任意时间间隔上, $y(t)$ 通过 $y=\eta(x)$ 近似于 $x(t)$. 使用这个隶属观点, 存在一些在任意时间上与 $x(t)$ 的近似值相关的结论. 特别地, 这些长时间近似结论是建立在构建不变流形的严格意义上的, 如第 8 章所描述的, 这个观点如下: 考虑方程

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= f(x,\eta(x)+z), \\ \frac{\mathrm{d}z}{\mathrm{d}t} &= \frac{1}{\varepsilon}g(x,\eta(x)+z)-\nabla\eta(x)f(x,\eta(x)+z).\end{aligned}$$

注意到 $y=\eta(x)+z$. 运用 $\nabla_y g(x,\eta(x))$ 是负定的事实, 可以证明 $z=\varepsilon\eta_1(x;\varepsilon)(\eta_1$ 在 $\varepsilon\to 0$ 上有界统一) 存在一个不变流形. x 和 z 的方程以初始条件 $z(0)=\varepsilon\eta_1(x(0);\varepsilon)$ 开始, 对于所有的检出时间将满足 $z(t)=\varepsilon\eta_1(x(t);\varepsilon)$. 此外, 当 $t\to\infty$ 时, 流形使得

$$|z(t)-\varepsilon\eta_1(x(t);\varepsilon)|\to 0,$$

从而可以得到 y 的一个不变流形, 它有如下形式:

$$y=\eta(x)+\varepsilon\eta_1(x;\varepsilon).$$

不变流形的存在性和唯一性可以通过各种方法证明，主要有 Lyapunov-Perron 法[136,314] 和 Hadamard图表转换法[330]. 这个领域中重要的工作是归于Fenichel[104,105]，他建立了一个相当广泛的正规双曲型不变流形的构造步骤. 文献 [57] 对证明不变流形存在性的 Lyapunov-Perron 法有一个清晰的说明.

图表和流的图表转化法的综述见文献 [305].

15.5 练　习

1. 在本章开始所陈述的对 g 的假设下, 说明对于任意 $t>0$, $\varphi_\varepsilon^t:\mathbb{R}^{d-l}\to\mathbb{R}^{d-l}$, (15.2.3) 是一个收缩映射, 并说明它的固定点是什么?

2. 去除对 η 和 f 的假设和完全有界, 并运用对 η 和 f 替换的线性增长假设, 证明一个类似于定理 15.2 的结论.

3. 对于 $0<\varepsilon\ll 1$ 和 $x\in\mathbb{R}^l$, $y\in\mathbb{R}^{d-l}$, 考虑方程

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t}&=Ax+\varepsilon f_0(x,y),\\ \frac{\mathrm{d}y}{\mathrm{d}t}&=-\frac{1}{\varepsilon}By+g_0(x,y).\end{aligned}\tag{15.5.1}$$

设 B 是对称正定的, 并且 $z(t,x_0;\eta)$ 是方程

$$\frac{\mathrm{d}z}{\mathrm{d}t}=Az+\varepsilon f_0(z,\eta(z)),\quad z(0,x_0;\eta)=x_0$$

的解, 给出 $\eta:\mathbb{R}^l\to\mathbb{R}^{d-l}$, 由

$$(T\eta)(x_0)=\int_{-\infty}^{0}\mathrm{e}^{-Bs/\varepsilon}g_0(z(s,x_0;\eta),\eta(z(s,x_0;\eta)))\mathrm{d}s$$

定义 $T\eta:\mathbb{R}^l\to\mathbb{R}^{d-l}$, 说明如果 η 是 T 的一个固定点, 则 $y=\eta(x)$ 是方程 (15.5.1) 的一个不变流形. (这被看成是解决不变流形结构的 Lyapunov-Perron 方法.)

4. 设 f_0, g_0 和所有的导数是一致有界的, 证明前一个问题中提到的 T 有一个固定点. (提示: 在 Lipschitz 图表中的一个空间上应用收缩映射变元是足够小且有足够小的 Lipschitz 常数.)

第 16 章　Markov 链的平均化：收敛性定理

16.1　引　　言

在本章中证明了一个与 Markov 链的平均化相关的结论, 所介绍的方法导致一个弱收敛型结论, 这个结论说明初始链条件下的期望与平均链条件下的期望近似. 这个方法是用来研究双 Markov 链的向后方程的. 基本假设 (5.2.2) 起着中心作用. 第 20 章中运用类似于本章的方法, 基于最大值原理证明了抛物型偏微分方程的均匀化理论. 主要定理的说明在 16.2 节, 其证明见 16.3 节; 16.4 节对参考文献进行了归纳总结.

16.2　定　　理

背景如第 9 章, 为了使证明更加明了, 主要针对有限状态空间. 设 $\mathcal{I}_x, \mathcal{I}_y \subseteq \{1, 2, \cdots\}$ 是有限集. 考虑 $\mathcal{I}_x \times \mathcal{I}_y$ 上的时间连续 Markov 链

$$z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

设这个 Markov 链是由 ε 参数化的, 并且其向后方程具有形式

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{1}{\varepsilon} Q_0 v + Q_1 v, \tag{16.2.1}$$

其中 Q_0, Q_1 由 (9.2.4) 给出. 设 $X(t)$ 是 $\mathcal{I}_x$ 上的一个 Markov 链, 并且具有向后方程

$$\frac{\mathrm{d}v_0}{\mathrm{d}t} = \overline{Q}_1 v_0, \tag{16.2.2}$$

其中 $\overline{Q}_1$ 由 (9.3.1) 给出. 感兴趣的是 $X(t)$ 的近似 $x(t)$. 注意到 Kolmogorov 方程暗含了近似过程, 这种近似过程的公式正好是第 9 章中运用形式渐近性质推导出来的.

注意到 $x(t)$ 不是自身的 Markov 过程, 只有 $(x(t), y(t))$ 是, 从而可以通过一个 Markov 随机过程来近似于一个非 Markov 随机过程. 更确切地说, 证明了在任意固定的时间下, $x(t)$ 的统计量近似于 $X(t)$ 的统计量. 也就是说, 可以证明在任意固定时间 t 上, $x(t)$ 弱收敛于 $X(t)$.

定理 16.1　对任意 $t > 0$, 当 $\varepsilon \to 0$, $x(t) \Rightarrow X(t)$.

16.3 证　明

设 v_0 如 (16.2.2) 中的定义, 则由构造可得

$$v_0 \in \mathcal{N}(Q_0), \quad \frac{\mathrm{d}v_0}{\mathrm{d}t} - Q_1 v_0 \perp \mathcal{N}(Q_0^{\mathrm{T}}),$$

因此, 存在 v_1, 使得

$$Q_0 v_0 = 0,$$

$$Q_0 v_1 = \frac{\mathrm{d}v_0}{\mathrm{d}t} - Q_1 v_0.$$

可以通过 v_1 与 Q_0^* 上的零空间正交使得 v_1 具有唯一性, 然而这个特别的选取也不是必要条件. 简单地要求选择一个有界解且在时间上具有有界导数.

对于任意这样的 v_1 以及对于由 (16.2.2) 给出的 v_0, 定义

$$r = v - v_0 - \varepsilon v_1.$$

将 $v = v_0 + \varepsilon v_1 + r$ 代入 (16.2.1), 并运用 v_0, v 的性质可得

$$\frac{\mathrm{d}v_0}{\mathrm{d}t} + \varepsilon\frac{\mathrm{d}v_1}{\mathrm{d}t} + \frac{\mathrm{d}r}{\mathrm{d}t} = \frac{1}{\varepsilon}Q_0 v_0 + Q_0 v_1 + \frac{1}{\varepsilon}Q_0 r + Q_1 v_0 + \varepsilon Q_1 v_1 + Q_1 r.$$

因此,

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \left(\frac{1}{\varepsilon}Q_0 + Q_1\right) r + \varepsilon q,$$

$$q = Q_1 v_1 - \frac{\mathrm{d}v_1}{\mathrm{d}t}.$$

接着$Q := \dfrac{1}{\varepsilon}Q_0 + Q_1$ 是一个 Markov 链的生成子, 因此, 用 $|\cdot|_\infty$ 表示在有限集 $\mathcal{I}_x \times \mathcal{I}_y$ 向量上的上确界范数, 并且由诱导算子范数可得

$$|\mathrm{e}^{Qt}|_\infty = 1. \tag{16.3.1}$$

因为 e^{Qt} 是一个随机矩阵, 因此, 可得 (5.2.2). 这个估计是非常值得注意的, 因为尽管 Q 在一个广义形式上取决于 ε, 而这个估计并不是.

通过常数变异公式可得

$$r(t) = \mathrm{e}^{Qt} r(0) + \varepsilon \int_0^t \mathrm{e}^{Q(t-s)} q(s)\mathrm{d}s. \tag{16.3.2}$$

对每一个 t, 将 $r(t)$, $q(t)$ 看成是 $\mathcal{I}_x \times \mathcal{I}_y$ 上的向量. 设 $v(i,k,0) = \phi(i)$, $v_0(i,0) = \phi(i)$, 则

$$v(i,k,t) = \mathbb{E}(\phi(x(t))|x(0) = i, y(0) = k), \tag{16.3.3a}$$

$$v_0(i,t) = \mathbb{E}(\phi(X(t))|X(0)=i). \tag{16.3.3b}$$

如果

$$v(i,k,t) \to v_0(i,t), \quad \varepsilon \to 0$$

对于任意 $\phi : \mathcal{I}_x \to \mathbb{R}$ 成立, 则可以证明对于固定的 t, $x(t)$ 弱收敛于 $X(t)$. (16.3.3) 暗含了 $r(0) = -\varepsilon v_1(0)$, 在 (16.3.2) 中运用 (16.3.1) 可得

$$\begin{aligned}
|r(t)|_\infty \leqslant & \varepsilon|\mathrm{e}^{Qt}|_\infty|v_1(0)|_\infty + \varepsilon\int_0^t |\mathrm{e}^{Q(t-s)}|_\infty|q(s)|_\infty \mathrm{d}s \\
\leqslant & \varepsilon|v_1(0)|_\infty + \varepsilon\int_0^t |q(s)|_\infty \mathrm{d}s \\
\leqslant & \varepsilon\left(|v_1(0)|_\infty + t\sup_{0\leqslant s\leqslant t}|q(s)|_\infty\right).
\end{aligned}$$

因此, 对于任意固定的 $t>0$, $r(t)\to 0(\varepsilon\to 0)$, 这遵循 $v\to v_0(\varepsilon\to 0)$. □

注 16.2 这个证明通常给出一个收敛率, 因为它说明对于 $0\leqslant t\leqslant T$ 有

$$|v(t)-v_0(t)|_\infty \leqslant C(T)\varepsilon.$$

16.4 讨论和参考

证明的结论仅说明了对于每一个固定的 t, 随机变量 $x(t)$ 收敛于 $X(t)$. 在 Skorokhod 拓扑中 (定义见文献 [120]), 证明弱收敛实际上发生在路径空间的这一更有趣结论也是可能的, 见文献 [291]; 推导 Markov 链的扩散极限也是可能的, 见文献 [182~184].

16.5 练 习

1. 考虑二态时间连续 Markov 链 y 且具有生成子

$$L = \frac{1}{\varepsilon}\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

和 $\mathcal{I}=\{-1,+1\}$ 态空间. 考虑 $\mathbb{T}^d$ 上的常微分方程由下式给出:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y),$$

其中 $f:\mathbb{T}^d\times\mathcal{I}\to\mathbb{R}^d$.

(a) 写出这个过程的生成子;

(b) 运用多尺度分析, 说明平均化随机微分方程是

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X),$$

其中

$$F(x) = \lambda f(x, +1) + (1 - \lambda) f(x, -1)$$

以及 $\lambda \in (0, 1)$ 必须是特定的.

2. 证明第 1 题中的断言：说明对于每一个固定的 $t \geqslant 0$ 都有 $x(t) \Rightarrow X(t)$.

3. 推测当 $F(x) = 0$ 并考虑大时间尺度时 x 的性质.

4. 验证第 9 章练习第 6 题推导的二态 Markov 链的近似值.

第 17 章　随机微分方程的平均化：收敛性定理

17.1　引　　言

本章的目的就是基于第 10 章中发展的随机微分方程的平均化原理来发展一个严格理论. 介绍 Itô 方程的运用, 并应用到一个仔细挑选的椭圆型偏微分方程 (**Poisson 方程**) 的解中, 是为了估计包含快变组分的积分 (见注 6.17). 这个估计方法的存在使随机微分方程的处理方法比纯常微分方程的处理方法明显得简单些; 在 17.4 节中给出了关于纯常微分方程的文献指南, 也考虑了平均化方程是常微分方程的情况, 这允许直接强收敛结论. 如果平均化自身是一个随机微分方程, 则弱收敛结论更自然. 这些弱收敛方法将在第 18 章中随机微分方程的均匀化背景下来说明.

17.2 节包含了一个定理介绍; 17.3 节是证明; 本章中提出的结论的推广及参考评论见 17.4 节.

17.2　定　　理

为了得到既简单又具有典型性的定理和说明, 在单位体 $\mathbb{T}^d$ 上研究问题

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y), \quad x(0) = x_0, \tag{17.2.1a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\varepsilon}g(x,y) + \frac{1}{\sqrt{\varepsilon}}\beta(x,y)\frac{\mathrm{d}V}{\mathrm{d}t}, \quad y(0) = y_0, \tag{17.2.1b}$$

其中 V 是 $\mathbb{R}^{d-l}$ 上的一个标准布朗运动, $f: \mathbb{T}^l \times \mathbb{T}^{d-l} \mapsto \mathbb{R}^l, g: \mathbb{T}^l \times \mathbb{T}^{d-l} \mapsto \mathbb{R}^{d-l}, \beta: \mathbb{T}^l \times \mathbb{T}^{d-l} \mapsto \mathbb{R}^{(d-l)\times(d-l)}$ 是光滑周期函数. 设 $B(x,y) = \beta(x,y)\beta(x,y)^{\mathrm{T}}$. 假定 $z = (x^{\mathrm{T}}, y^{\mathrm{T}})^{\mathrm{T}}$ 时有

$$\exists \bar{\beta} > 0: \ \langle \xi, B(x,y)\xi \rangle \geqslant \bar{\beta}|\xi|^2, \forall \xi \in \mathbb{R}^{d-l}, z \in \mathbb{T}^d. \tag{17.2.2}$$

也设初始条件是确定的, 这对平均化结论并不是必要条件, 而只是为了简化的目的.

回顾假设 (17.2.2), 则在 (17.2.1b) 中固定 $x = \xi$ 得到的过程具有遍历性 (定理 6.16). 这样期望在这个遍历性过程的不变测度下, 通过平均化 f 得到一个 x 进化的有效方程. 接下来让这种思想更确切.

由 (10.5.1) 给出的过程 φ_ξ^t 具有遍历性且有一个光滑不变密度 $\rho^\infty(y;\xi)$. 当 $x=\xi$ 时进行评估, 这种不变密度张成了 $\mathcal{L}_0$ 的伴随 $\mathcal{L}_0^*$ 的零空间, 其中 $\mathcal{L}_0$ 是由下式给出:

$$\mathcal{L}_0 = g(x,y)\cdot\nabla_y + \frac{1}{2}B(x,y):\nabla_y\nabla_y. \tag{17.2.3}$$

$\mathcal{L}_0$ 和 $\mathcal{L}_0^*$ 都具有周期边界条件, 则平均化方程为

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X), \tag{17.2.4a}$$

$$F(\xi) = \int_{\mathbb{T}^{d-l}} f(\xi,y)\rho^\infty(y;\xi)\mathrm{d}y. \tag{17.2.4b}$$

注意到由构造知 $F:\mathbb{T}^l\mapsto\mathbb{R}^l$ 是周期性的, 这些结论形式在第 10 章中确切地提出, 尤其是对特殊流和扩散系数的专门研究见 (17.2.1).

定理 17.1　设 $p>1$ 和 $X(0)=x_0$, 则 (17.2.1) 的解函数 $x(t)$ 收敛于 (17.2.4) 的解函数 $X(t)$, 在 $L^p(\Omega, C([0,T],\mathbb{T}^l))$ 中, 对任意 $T>0$, 存在 $C=C(T)$, 使得

$$\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}|x(t)-X(t)|^p\right)\leqslant C\varepsilon^{p/2}.$$

17.3　证　明

回顾 $\mathcal{L}_0$ 是 $\varphi_x^t(y)$ 的生成子, 其中 $\mathcal{L}_0$ 被看成是固定参量, 由 (17.2.3) 给出. 从而 $\mathcal{L}_0$ 仅是 y 中的一个微分算子, x 是一个参量. 接下来, 设 $\phi(x,y)$ 是如下椭圆边值问题的解:

$$\begin{aligned}&\mathcal{L}_0\phi(x,y) = f(x,y)-F(x),\\ &\int_{\mathbb{T}^{d-l}}\phi(x,y)\rho^\infty(y;x)\mathrm{d}y = 0,\\ &\phi(x,\cdot)\text{在}\mathbb{T}^{d-l}\text{上具有周期性}.\end{aligned}$$

这被称为 Poisson 方程. 由构造

$$\int_{\mathbb{T}^{d-l}}(f(x,y)-F(x))\rho^\infty(y;x)\mathrm{d}y = 0$$

及 ρ^∞ 张成了 $\mathcal{N}(\mathcal{L}_0^*)$. 因此, 由 Fredholm 性质或定理 7.8 知 ϕ 有唯一解.

引理 17.2　函数 f, ϕ, $\nabla_x\phi$, $\nabla_y\phi$ 和 β 光滑且有界.

证明　f 和 β 的性质遵从这样一个事实, 它们是定义在单位体上的, 并且由假设知其具有各阶次导数. 由于不变密度 ρ^∞ 是圆环上一个椭圆特征值问题的解, 也是光滑的和周期的, 因此, F 是光滑的和周期的, 从而 $f-F$ 是光滑的和周期的, 那么 ϕ 和它所有的导数也是光滑的和周期的. □

定理 17.1 的证明 注意到 (17.2.1) 的生成子是

$$\mathcal{L} = \frac{1}{\varepsilon}\mathcal{L}_0 + \mathcal{L}_1,$$

其中 $\mathcal{L}_0$ 由 (17.2.3) 给出, 并且

$$\mathcal{L}_1 = f(x,y)\cdot\nabla_x.$$

接下来, 将 Itô 方程 (引理 6.5) 应用到 $\phi(x(t),y(t))$, 并在时间上积分, 可得如下表达式:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(x,y) = \frac{1}{\varepsilon}(\mathcal{L}_0\phi)(x,y) + f(x,y)\cdot\nabla_x\phi(x,y) + \frac{1}{\sqrt{\varepsilon}}\nabla_y\phi(x,y)\beta(x,y)\frac{\mathrm{d}V}{\mathrm{d}t}.$$

由 $\mathcal{L}_0\phi = f - F$ 可得

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} =& F(x) + (\mathcal{L}_0\phi)(x,y)\\ =& F(x) + \varepsilon\frac{\mathrm{d}\phi}{\mathrm{d}t} - \varepsilon f(x,y)\cdot\nabla_x\phi(x,y) - \sqrt{\varepsilon}\nabla_y\phi(x,y)\beta(x,y)\frac{\mathrm{d}V}{\mathrm{d}t}\end{aligned} \tag{17.3.1}$$

(再一次在时间上积分使得一个正式的表达更严格). 函数 f, ϕ 和 $\nabla_x\phi$ 都是光滑的, 并且由定理 17.2 知它们有界. 因此, 存在一个常数 $C>0$, 使得

$$\theta(t) := (\phi(x(t),y(t)) - \phi(x(0),y(0))) - \int_0^t f(x(s),y(s))\cdot\nabla_x\phi(x(s),y(s))\mathrm{d}s$$

满足

$$\sup_{0\leqslant t\leqslant T}|\theta(t)| \leqslant C.$$

接下来, 考虑鞅

$$M(t) := -\int_0^t \nabla_y\phi(x(s),y(s))\beta(x(s),y(s))\mathrm{d}V(s).$$

由于 $\nabla_y\phi$, β 光滑且有界, 由定理 17.2, Itô 等距给出

$$\begin{aligned}\mathbb{E}|\langle M\rangle_t|^2 \leqslant& C\int_0^t \mathbb{E}|\nabla_y\phi(x(s),y(s))\beta(x(s),y(s))|_F^2\mathrm{d}s\\ \leqslant& Ct.\end{aligned}$$

类似地 (见 (3.4.8)), 对于 $p\geqslant 1$,

$$\mathbb{E}|\langle M\rangle_t|^{p/2} \leqslant C. \tag{17.3.2}$$

(17.3.1) 的严格表达式为

$$x(t) = x(0) + \int_0^t F(x(s))\mathrm{d}s + \varepsilon\theta(s) + \sqrt{\varepsilon}M(t).$$

从 (17.2.4) 也可知

$$X(t) = X(0) + \int_0^t F(X(s))\mathrm{d}s.$$

设 $e(t) = x(t) - X(t)$, 运用 $e(0) = 0$, 使得

$$e(t) = \int_0^t (F(x(s)) - F(X(s)))\mathrm{d}s + \varepsilon\theta(t) + \sqrt{\varepsilon}M(t).$$

由于 F 在 $\mathbb{T}^l$ 上是 Lipschitz 的, 则对于 $t \in [0,T]$ 可得

$$|e(t)| \leqslant \int_0^t L|e(s)|\mathrm{d}s + \varepsilon C + \sqrt{\varepsilon}|M(t)|.$$

因此, 由 (17.3.2) 和 Burkholder-Davis-Gundy 不等式及定理 3.22 可得①

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}|e(t)|^p\right) &\leqslant C\left(\varepsilon^p + \varepsilon^{p/2}\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}|M(t)|^p\right) + L^pT^{p-1}\int_0^T \mathbb{E}|e(s)|^p\mathrm{d}s\right)\\ &\leqslant C\left(\varepsilon^p + \varepsilon^{p/2}\mathbb{E}(|\langle M\rangle_T|)^{p/2} + L^pT^{p-1}\int_0^T \mathbb{E}|e(s)|^p\mathrm{d}s\right)\\ &\leqslant C\left(\varepsilon^{p/2} + \int_0^T \mathbb{E}\sup_{0\leqslant\tau\leqslant s}|e(\tau)|^p\mathrm{d}s\right).\end{aligned}$$

由引理 4.4 中 Gronwall 不等式的积分可推导出

$$\mathbb{E}\left(\sup_{0\leqslant t\leqslant T}|e(t)|^p\right) \leqslant C\varepsilon^{p/2}.$$

这样定理也就被证明出来了. □

17.4 讨论和参考

收敛结论证明了强收敛: 将 (x,y) 随机微分方程的每条路径和 X 常微分方程的近似值相比较. 这个强收敛结论是可能的, 主要是因为在这种情况下极限近似是确定的. 正如 (10.7.1) 中提出的, 当这个近似本身是随机时, 则会更自然地研究弱收敛型结论 (见练习第 1 题), 对于后一类型的结论见 [94].

① 对于 C 是一个常数, 独立于 ε, 随着方程推导而改变.

在这个证明中, ϕ 的 Poisson 方程和 Itô 方程在 ϕ 中的应用起到关键作用. 随机微分方程平均化理论的证明见 [111, 第 7 章], 但是注意到在其中并没有对扰动 Poisson 方程进行系统的使用.

常微分方程的平均化结论是由某些不同的方法来证明的, 因为基于 Itô 公式的方法并不适用, 主要因为 Poisson 方程不再是椭圆的. 很多研究常微分方程平均化的最初动机是源于扰动可积 Hamiltonian 系统的平均化, 见文献 [135].

17.5 练　　习

1. 当 $\alpha(x,y)\equiv 1$ 时考虑 (10.7.1), 写出 X 的平均化动态性以及本章证明的从 x 到 X 强收敛的方法修正.

2. 当 $d=2$, $l=1$ 且 $\alpha(x,y)\equiv y^2$ 时考虑 (10.7.1), 如果 y 是一个标量 OU 过程 (6.4.4), 独立于 x, 写出 X 的平均化动态性, 并运用 OU 过程的性质证明弱收敛.

3. 考虑快 / 慢系统

$$\dot{x}=f(x)\eta^2, \quad \dot{\eta}=-\frac{1}{\varepsilon}\eta+\frac{1}{\sqrt{\varepsilon}}\dot{W},$$

其中 $W(t)$ 是一个一维标准 Brownian 运动. 运用函数 $f(x)\eta^2$ 的 Itô 公式证明一个强收敛定理, 并说明它与基于本章中使用的 Poisson 方程的方法之间的关系.

第 18 章　随机微分方程的均匀化：收敛性定理

18.1　引　　言

本章在第 11 章推导的均匀化原理的基础上发展了严格理论. 为了阐明这个原理的思想, 考虑一个简单的背景, 即整个问题是在单位体上提出的. 此外, 在 Skim-product 背景下研究, 其中快速过程是独立于慢过程的. 最后假设慢过程中的波动项仅取决于快速过程. 这些假设考虑到一个简易的证明, 它包含了这个主要思想的精髓. 如前面章节中, 适当的 Poisson 方程和 Itô 格式的使用起着中心作用 (见注 6.17); 18.2 节包含了定理说明; 18.3 节是定理的证明; 本章还证明了收敛性定理的推广以及一些文献参考, 在 18.4 节给出.

18.2　定　　理

在 $\mathbb{T}^d$ 上考虑常微分方程的如下系统:

$$
\left\{
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{1}{\varepsilon}f_0(y) + f_1(x,y) + \alpha\frac{\mathrm{d}U}{\mathrm{d}t}, & (18.2.1\mathrm{a})\\
\frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{1}{\varepsilon^2}g(y) + \frac{1}{\varepsilon}\beta(y)\frac{\mathrm{d}V}{\mathrm{d}t}, & (18.2.1\mathrm{b})
\end{aligned}
\right.
$$

其中 U(或 V) 是 $\mathbb{R}^l$(或 $\mathbb{R}^{d-l}$) 上的标准 Brownian 运动, 并且 $\alpha\in\mathbb{R}^{l\times l}$, 是一个常数矩阵. 这两个 Brownian 运动假设是相互独立的. 函数 $f_0:\mathbb{T}^{d-l}\mapsto\mathbb{R}^l, f_1:\mathbb{T}^l\times\mathbb{T}^{d-l}\mapsto\mathbb{R}^l, g:\mathbb{T}^{d-l}\mapsto\mathbb{R}^{d-l}, \beta:\mathbb{T}^{d-l}\mapsto\mathbb{R}^{(d-l)(d-l)}$ 都是光滑和周期性的. 设 $B(y)=\beta(y)\beta(y)^{\mathrm{T}}$, 也设

$$
\exists\overline{\beta}>0, \langle\xi,B(y)\xi\rangle\geqslant\overline{\beta}|\xi|^2,\quad \forall\xi\in\mathbb{R}^{d-l}, y\in\mathbb{T}^{d-l}.
$$

回顾在这个假设下, (18.2.21b) 中的过程 y 具有遍历性 (见结论 6.16), 从而通过在这个遍历性过程的不变测度上平均化 f_1, 以及通过检验由 f_0 诱发的波动, 可以得到 x 演变的有效方程. 本章的目的就是使这些思想更加严密化.

过程 y 的生成子由下式给出:

$$
\mathcal{L}_0 = g(y)\cdot\nabla_y + \frac{1}{2}B(y):\nabla_y\nabla_y,
$$

并且上式具有周期边界条件. 在这个例子中, 快速过程是独立于 x 的, 对 $\mathcal{L}_0$ 来说, 方程 (17.2.3) 是很简单的, 当消除了 ξ 依赖时, 则 $\varphi^t(y)$ 由 (10.5.1) 给出. 过程 $\varphi^t(y)$ 具有遍历性且有一个光滑不变密度 $\rho^\infty(y)$, 这个函数张成 $\mathcal{L}_0^*$ 的零空间. 注意到 $\varphi^t(\cdot)$ 和 $\rho^\infty(\cdot)$ 独立于这里的 ξ, 因为 g, β 仅取决于 y. 设

$$\int_{\mathbb{T}^{d-l}} f_0(y)\rho^\infty(y)\mathrm{d}y = 0.$$

在这种假设下, 运用第 11 章中的计算, 并且均匀化随机微分方程 (18.2.1) 可得

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) + A\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{18.2.2}$$

其中

$$F(\xi) = \int_{\mathbb{T}^{d-l}} f_1(\xi, y)\rho^\infty(y)\mathrm{d}y$$

和

$$AA^{\mathrm{T}} = \alpha\alpha^{\mathrm{T}} + \int_{\mathbb{T}^{d-l}} (f_0(y)\otimes\Phi(y) + \Phi(y)\otimes f_0(y))\rho^\infty(y)\mathrm{d}y.$$

回顾 $\Phi(y)$ 是如下单元问题的解:

$$\mathcal{L}_0\Phi(y) = -f_0(y), \tag{18.2.3}$$

$$\int_{\mathbb{T}^{d-l}} \Phi(y)\rho^\infty(y)\mathrm{d}y = 0, \quad \Phi(y)\text{在}\mathbb{T}^{d-l}\text{上是具有周期性}.$$

将 Fredholm 性质应用到具有周期边界条件的椭圆型偏微分方程可知这个问题有唯一解, 见 7.2.3 小节. 第 11 章准确地给出了 A 和 F 的具体形式, 同时本章中也对特殊流及分散系数进行了专门研究.

注意到对于 $\xi\in\mathbb{R}^l$, $\phi = \Phi\cdot\xi$, 定理 11.3 的证明说明了 AA^{T} 是正定的, 并且

$$\langle\xi, AA^{\mathrm{T}}\xi^{\mathrm{T}}\rangle = |\alpha^{\mathrm{T}}\xi|^2 + \int_{\mathbb{T}^{d-l}} |\beta^{\mathrm{T}}(y)\nabla\phi(y)|^2\rho^\infty(y)\mathrm{d}y.$$

因此, X 的随机微分方程是适定的. 注 11.4 表明了

$$AA^{\mathrm{T}} = \alpha\alpha^{\mathrm{T}} + \alpha_2\alpha_2^{\mathrm{T}},$$

$$\alpha_2\alpha_2^{\mathrm{T}} = \int_{\mathbb{T}^{d-l}} \rho^\infty(y)(\nabla\Phi(y)\beta(y)\otimes\nabla\Phi(y)\beta(y))\mathrm{d}y. \tag{18.2.4}$$

基于本章用到的证明, 便会很自然地在 Poisson 方程中形成有效扩散矩阵的这种形式.

定理 18.1 设 $x(t)$ 是 (18.2.1) 的解, $X(t)$ 是 (18.2.2) 的解, 并且 $X(0) = x(0)$, 则对于任意 $T > 0$, 在 $C([0,T],\mathbb{T}^l)$ 上 $x\Rightarrow X$.

18.3 证　明

证明的构造如下: 说明过程 $x(t)$ 满足

$$x(t) = x(0) + \int_0^t F(x(s))\mathrm{d}s + \alpha U(t) + M_2(t) + \eta(t). \tag{18.3.1}$$

期望极限过程满足

$$X(t) = x(0) + \int_0^t F(X(s))\mathrm{d}s + \alpha U(t) + \alpha_2 W(t),$$

其中扩散系数 α_2 的定义在 (18.2.4) 中. 在 x 的方程中, 这一项表示从期望极限过程中得到的偏差在 $C([0,T],\mathbb{R}^{3l})$ 上满足

$$(\alpha U, \eta, M_2) \Rightarrow (\alpha U, 0, \alpha_2 W), \tag{18.3.2}$$

则通过概率测度的弱收敛在连续映射条件下保持不变这样一个事实, 运用如下引理推导期望收敛.

引理 18.2 设 $w \in C([0,T],\mathbb{R}^r)$①, $F \in C^1(\mathbb{T}^l,\mathbb{T}^l)$ 以及 $D \in \mathbb{R}^{l\times r}$, 存在唯一的 $u \in C([0,T],\mathbb{T}^l)$ 满足积分方程

$$u(t) = u(0) + \int_0^t F(u(s))\mathrm{d}s + Dw(t),$$

并且映射 $w \mapsto u$ 是一个从 $C([0,T],\mathbb{R}^r)$ 到 $C([0,T],\mathbb{T}^l)$ 的连续映射.

证明 存在性和唯一性遵从基于如下迭代的一个标准压缩映射:

$$u^{n+1}(t) = u(0) + \int_0^t F(u^n(s))\mathrm{d}s + Dw(t),$$

其中 F 是一个含有常数 L 的全局 Lipschitz 函数. 为了连续性, 考虑方程

$$u^{(i)}(t) = u(0) + \int_0^t F(u^{(i)}(s))\mathrm{d}s + Dw^{(i)}(t),$$

其中 $i = 1, 2$. 设 $e = u^1 - u^2, \delta = w^1 - w^2$, 代入可得

$$e(t) = \int_0^t (F(u^1(s)) - F(u^2(s)))\mathrm{d}s + D\delta(s).$$

因此,

$$|e(t)| \leqslant \int_0^t L|e(s)|\mathrm{d}s + |D||\delta(t)|.$$

① 这里的 $w(t)$ 是一个任意连续轨迹, 不一定必须是 Brownian 轨迹.

从引理 4.4 中 Gronwall 不等式的积分形式可知, 上式遵从

$$\sup_{0\leqslant t\leqslant T}|e(s)|\leqslant C\sup_{0\leqslant t\leqslant T}|\delta(t)|.$$

这样就建立了连续性. □

在定理 18.1 的建立中, 将两次用到 Itô 格式：一次是对 $\Phi(y(s))$, 其中 Φ 是单胞问题 (18.2.3) 的解, 另一次是对 $\chi(x(s),y(s))$, 其中 χ 是如下 Poisson 方程的解：

$$\mathcal{L}_0\chi(x,y)=f_1(x,y)-F(x), \tag{18.3.3}$$

$$\int_{\mathbb{T}^{d-l}}\chi(x,y)\rho^{\infty}(y)\mathrm{d}y=0,\quad \chi(x,y)\text{在}\mathbb{T}^{d-l}\text{上具有周期性}.$$

由于在 $\mathbb{T}^{d-l}$ 上 f_1-F 的平均值为零, 则由 Fredholm 性质或定理 7.8 可知式 (18.3.3) 有唯一解. 如下引理的证明与引理 17.2 的证明非常相似, 因此, 在这省略了.

引理 18.3 函数 f_0,f_1,Φ,χ 和它们所有的导数都是光滑且有界的.

定理 18.1 的证明 设 $\chi=\chi(x(t),y(t))$, 其中 $(x(t),y(t))$ 是 (18.2.1) 的解. 注意到由 Itô 公式, 引理 6.5 可得

$$\frac{\mathrm{d}\chi}{\mathrm{d}t}=\frac{1}{\varepsilon^2}\mathcal{L}_0\chi+\frac{1}{\varepsilon}\mathcal{L}_1\chi+\mathcal{L}_2\chi+\nabla_x\chi\alpha\frac{\mathrm{d}U}{\mathrm{d}t}+\frac{1}{\varepsilon}\nabla_y\chi\beta\frac{\mathrm{d}V}{\mathrm{d}t}. \tag{18.3.4}$$

严格来说, 应该是积分形式. 因此, 算子 $\mathcal{L}_0$ 正如之前定义的一样, 并且 $\mathcal{L}_1$, $\mathcal{L}_2$ 是由

$$\mathcal{L}_1=f_0(y)\cdot\nabla_x,\quad \mathcal{L}_2=f_1(x,y)\cdot\nabla_x+\frac{1}{2}\alpha\alpha^{\mathrm{T}}:\nabla_x\nabla_x$$

给出的. 接下来定义

$$\begin{aligned}\theta(t)=&\varepsilon^2(\chi(x(t),y(t))-\chi(x(0),y(0)))\\&-\varepsilon\int_0^t(\mathcal{L}_1\chi)(x(s),y(s))\mathrm{d}s-\varepsilon^2\int_0^t(\mathcal{L}_2\chi)(x(s),y(s))\mathrm{d}s\end{aligned} \tag{18.3.5}$$

和

$$\begin{aligned}M_1(t)=&\varepsilon^2\int_0^t\nabla_x\chi(x(s),y(s))\alpha\mathrm{d}U(s)\\&+\varepsilon\int_0^t\nabla_y\chi(x(s),y(s))\beta(y(s))\mathrm{d}V(s).\end{aligned} \tag{18.3.6}$$

由于 χ,f_0 和 f_1 以及它们所有的导数都是有界的, 可得

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}|\theta(t)|^p\leqslant C\varepsilon^p. \tag{18.3.7}$$

此外,

$$\begin{aligned}\mathbb{E}|\langle M_1\rangle|^2 =&\varepsilon^4\int_0^t|\nabla_x\chi(x(s),y(s))\alpha|_F^2\mathrm{d}s\\&+\varepsilon^2\int_0^t|\nabla_y\chi(x(s),y(s))\beta(y(s))|_F^2\mathrm{d}s.\end{aligned}$$

更一般地, 对于每一个 $p\geqslant 1$ 有

$$\begin{aligned}\mathbb{E}|\langle m_1\rangle|^{p/2} \leqslant&C_1\varepsilon^{2p}\int_0^t|\nabla_x\chi(x(s),y(s))\alpha|_F^p\mathrm{d}s\\&+C_2\varepsilon^p\int_0^t|\nabla_y\chi(x(s),y(s))\beta(y(s))|_F^p\mathrm{d}s.\end{aligned}$$

由定理 3.22 的 Burkholder-Davis-Gundy 不等式和 (3.4.8) 可以推导出

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}|M_1(t)|^p\leqslant C\varepsilon^p. \tag{18.3.8}$$

因此, 由 (18.3.4) 和 (18.3.3) 可推导出

$$\int_0^t(f_1(x(s),y(s))-F(x(s)))\mathrm{d}s=r(t):=\theta(t)-M_1(t), \tag{18.3.9}$$

其中已经说明了, 在 $L^p(\Omega,C([0,T],\mathbb{R}^l))$ 上,

$$r=O(\varepsilon). \tag{18.3.10}$$

由于 $\varPhi$ 是独立于 x 的, 接下来将 Itô 格式应用到 $\varPhi$ 中求解 (18.2.3) 可得

$$\frac{\mathrm{d}\varPhi}{\mathrm{d}t}=\frac{1}{\varepsilon^2}\mathcal{L}_0\varPhi+\frac{1}{\varepsilon}\nabla_y\varPhi\beta\frac{\mathrm{d}V}{\mathrm{d}t}$$

(和平常一样严格来说应该是积分形式). 因此,

$$\frac{1}{\varepsilon}\int_0^t f_0(y(s))\mathrm{d}s=\varepsilon(\varPhi(y(0))-\varPhi(y(s)))+\int_0^t\nabla_y\Phi(y(s))\beta(y(s))\mathrm{d}V(s). \tag{18.3.11}$$

由于 $\varPhi$ 在 $\mathbb{T}^{d-l}$ 上有界, 当 $\varepsilon\to 0$ 时, 在 $L^p(\Omega,C([0,T],\mathbb{R}^l))$ 上可推导出

$$\varepsilon(\varPhi(y(\cdot))-\varPhi(y(0)))=O(\varepsilon). \tag{18.3.12}$$

设

$$M_2(t)=\int_0^t(\nabla_y\varPhi)(y(s))\beta(y(s))\mathrm{d}V(s),$$

注意到 $y(s)=\varphi^{s/\varepsilon^2}(y)$, 其中 $y(0)=y$. 重新标注时间尺度, 并运用注 6.3 Brownian 运动尺度的性质, 依法则可得密度

$$M_2(t) = \varepsilon \int_0^{t/\varepsilon^2} (\nabla_y \Phi)(\varphi^\tau(y))\beta(\varphi^\tau(y))\mathrm{d}V(\tau).$$

Martingale 中心极限定理 3.33 暗含了当 $\varepsilon \to 0$ 时, 在 $C([0,T],\mathbb{R}^l)$ 上有

$$M_2 \Rightarrow \alpha_2 W, \tag{18.3.13}$$

其中 $W(t)$ 是标准 Brownian 运动, 并且

$$\begin{aligned}\alpha_2\alpha_2^{\mathrm{T}} &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \nabla_y\Phi(\varphi^s(y))\beta(\varphi^s(y))\otimes\nabla_y\Phi(\varphi^s(y))\beta(\varphi^s(y))\mathrm{d}s\\ &= \int_{\mathbb{T}^{d-l}}\rho^\infty(y)(\nabla_y\Phi(y)\beta(y)\otimes\nabla_y\Phi(y)\beta(y))\mathrm{d}y.\end{aligned}$$

在 (18.2.1a) 中结合 (18.3.9), (18.3.11) 可得 (18.3.1), 其中

$$\eta(t) = r(t) + \varepsilon(\Phi(y(0)) - \Phi(y(t))).$$

在 $C([0,T],\mathbb{R}^{2l})$ 上有

$$(\alpha U, M_2) \Rightarrow (\alpha U, \alpha_2 W).$$

因为 U 和 M_2 是独立的, 由 (18.3.10), (18.3.12) 可得在 $L^p(\Omega, C([0,T],\mathbb{R}^l))$ 上有

$$\eta \to 0.$$

因此, 由定理 3.30 知 (18.3.2) 成立, 即在 $C([0,T],\mathbb{R}^{3l})$ 上有

$$(\alpha U, \eta, M_2) \Rightarrow (\alpha U, 0, \alpha_2 W).$$

由定理 18.2 可知, (18.3.1) 中从 $C([0,T],\mathbb{R}^{3l})$ 到 $C([0,T],\mathbb{T}^l)$ 的映射 $(U,\eta,M_2)\to x$ 是连续的. 因此, 在连续映射下, 弱收敛是保守的 (见定理 3.28), 在 $C([0,T],\mathbb{T}^l)$ 上可推导出 $x(t)\Rightarrow X(t)$, 其中 X 是如下方程的解:

$$X(t) = x(0) + \int_0^t F(X(s))\mathrm{d}s + \alpha U(t) + \alpha_2 W(t).$$

18.4 讨论和参考

本章中考虑到的类型, 即随机微分方程奇异扰动系统的极限定理, 早在 19 世纪 60 年代就已经开始研究了, 如见文献 [166, 300, 301]. 这个理论更深一步的发展是在 19 世纪 70 年代, 见文献 [33; 94, ch.12, 243; 291] 以及其中的参考文献.

通过假设 X 的极限随机微分方程已被附加噪声, 这里提出的证明已经过简化, 从而可以直接使用鞅中心极限定理. 一般情况下, 极限过程有与普通状态相关的噪

声, 这使得收敛性的证明更复杂, 并且要用到随机微分方程解的鞅方程, 见文献 [94, ch.12].

对于 Poisson 方程在随机微分方程极限定理证明中所起到的基础性作用, 人们已经认知到很久了[181,240,241,246]. 适当的 Poisson 方程在非紧状态空间下的奇异扰动随机微分方程极限定理 (扩散近似) 中的系统应用, 最近已经由 Pardoux 和 Veretennikov 在一系列的论文中实行[249~251].

通过介绍附加 Poisson 方程以及 Itô 格式在它们的解中的应用, 则结合 Dambis-Dubins-Schwarz 定理, 获得收敛率也是可能的, 见 3.6 节和文献 [134]. 这些误差估计在具有多尺度构造的随机微分方程参数估计的研究中起到重要作用, 见文献 [258]. 关于鞅的一个相当一般的强收敛定理的证明见文献 [40].

18.5 练 习

1. 考虑如下标量随机微分方程组的耦合对:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,y) - \frac{1}{\varepsilon}y + (y^2 + a(x))\frac{\mathrm{d}U}{\mathrm{d}t}, \tag{18.5.1a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{1}{\varepsilon^2}y + \frac{1}{\varepsilon}\sqrt{2}\frac{\mathrm{d}V}{\mathrm{d}t}. \tag{18.5.1b}$$

(a) 写出均匀化方程;

(b) 设 f,a 和所有的导数有界, 运用 OU 过程的精确解猜测一个均匀化方程, 并证明其收敛性定理.

2. (a) 考虑快/慢系统

$$\dot{x} = \frac{f(x)y}{\varepsilon}, \quad \dot{y} = -\frac{\alpha}{\varepsilon^2}y + \sqrt{\frac{2\lambda}{\varepsilon^2}}\dot{\beta};$$

(b) 通过将 Itô 格式应用到 $x(t)$ 和 $y(t)$ 适当的函数中来证明均匀化方程是

$$\dot{X} = \frac{\lambda}{\alpha^2}f(X)f'(X) + \sqrt{\frac{2\lambda}{\alpha^2}}f(X)\dot{\beta},$$

并运用这个方法得到误差估计;

(c) 设快过程是稳态的. 运用 Itô 公式证明一个强 (顺时) 极限定理. (提示: 需要运用关于 Ornstein-Uhlenbeck 过程的轨迹估计.)

3. 考虑本章一维情况下研究的问题. 结合这里提出的对 Dambis-Duchwarz 定理的分析证明一个强 (顺时) 收敛定理.

4. 考虑偏微分方程 (18.2.1), 其中 $x \in \mathbb{R}^l$, $y \in \mathbb{R}^{d-l}$. 仔细说明如下性质: 在这种情况下, Poisson 方程的解有这样一种特性, 即可以证明均匀化定理.

第 19 章　椭圆型偏微分方程的均匀化：收敛性定理

19.1　引　　言

本章中证明了具有周期系数和 Dirichlet 边界条件的二阶一致椭圆型偏微分方程的两个均匀化定理. 证明的方法就是运用双尺度收敛. 这个方法为推导均匀化方程提供了一个高雅的方式, 并且这个单元问题需要通过一个称为双尺度系统的耦合系统方程来定义. 双尺度收敛的方法也适用于运输方程, 并且将在第 21 章的内容中用到. 19.2 节对两个定理进行了说明; 接下来的两节是它们的证明; 本章以评论和参考文献结束.

19.2　定　　理

定理 19.1　设 u^ε 是如下方程的弱解:

$$-\nabla \cdot (A^\varepsilon \nabla u^\varepsilon) = f, \quad x \in \Omega, \tag{19.2.1a}$$

$$u^\varepsilon = 0, \quad x \in \partial\Omega, \tag{19.2.1b}$$

其中 $f = f(x) \in L^2(\Omega), \Omega \subset \mathbb{R}^d$ 有界, 并且 $A^\varepsilon = A\left(\dfrac{x}{\varepsilon}\right)$, $A \in M_{\mathrm{per}}(\alpha, \beta, \mathbb{T}^d)$, $0 < \alpha \leqslant \beta < \infty$. 此外, 设 u 是如下均匀化问题的弱解:

$$-\nabla \cdot (\bar{A} \nabla u) = f, \quad x \in \Omega, \tag{19.2.2a}$$

$$u = 0, \quad x \in \partial\Omega, \tag{19.2.2b}$$

其中 $\overline{A}$ 是由

$$\overline{A} = \int_{\mathbb{T}^d} (A(y) + A(y)\nabla_y \chi(y)^{\mathrm{T}})\mathrm{d}y \tag{19.2.3}$$

给出的, 并且向量场 $\chi(y)$ 是如下单胞问题的弱解:

$$-\nabla_y \cdot (\nabla_y \chi A^{\mathrm{T}}) = \nabla_y \cdot A^{\mathrm{T}}, \quad \chi(y)\text{是 1 周期}, \tag{19.2.4}$$

则在 $H_0^1(\Omega)$ 上有如下弱收敛:

$$u^\varepsilon \rightharpoonup u,$$

并且在 $L^2(\Omega)$ 上有如下强收敛：

$$u^\varepsilon \to u.$$

除了基本的均匀化定理外, 还将证明一个结论, 这个结论说明了双尺度展开式中保留的附加项的确给出了改善的近似值. 如下结论说明只要考虑一阶校正场就可以得到 $H_0^1(\Omega)$ 上的强收敛:

定理 19.2 如定理 19.1 中考虑 $u^\varepsilon(x)$ 和 $u(x)$. 设 $f \in L^2(\Omega)$, 其中 $\partial\Omega$ 充分光滑以使得 $u \in H^2(\Omega) \cap H_0^1(\Omega)$, 并且系数矩阵 A 使得单胞问题 (19.2.4) 具有解 $\chi \in C_{\mathrm{per}}^1(\mathbb{T}^d)$, 则

$$\lim_{\varepsilon \to 0} \left\| u^\varepsilon(x) - \left(u(x) + \varepsilon \chi\left(\frac{x}{\varepsilon}\right) \cdot \nabla u(x) \right) \right\|_{H^1(\Omega)} = 0. \tag{19.2.5}$$

注 19.3 如果 $\partial\Omega$ 是一个超曲面或者凸多面体, 则满足关于它的光滑性假设. 例如, 为了使 $\chi \in C_{\mathrm{per}}^1(\mathbb{T}^d)$, $A(y)$ 不得不充分正则, $A(y) \in C_{\mathrm{per}}^1(\mathbb{T}^d; \mathbb{R}^{d\times d})$ 更是绰绰有余.

19.3 证明：L^2 上的强收敛

本节运用双尺度收敛方法证明了均匀化定理及定理 19.1. 证明开始之前, 作了一些关于这个方法的备注. 分析的第 1 步是运用第 7 章的能量估计推导 u^ε 和 ∇u^ε 具有双尺度收敛子序列, 这个子序列是通过一对函数 $\{u(x), u_1(x,y)\}$ 定义的. 第 2 步是为了表征双尺度极限, 在 (19.2.2) 弱解的定义中使用一个具有如下形式的测度函数:

$$\phi^\varepsilon(x) = \phi_0(x) + \varepsilon \phi_1\left(x, \frac{x}{\varepsilon}\right), \tag{19.3.1}$$

以这种方式, 可以获得出现在第 1 步中分析的函数的双系统方程, 即 $\{u, u_1\}$; 称这些方程为双尺度系统, 见引理 19.5. 第 3 步就是在引理 19.6 中, 运用 Lax-Milgram 定理证明这个系统的存在性和唯一性. 最后一步, 引理 19.7, 就是运用变量分离解耦方程的这个系统, 说明它引起的均匀化方程 (19.2.2)~(19.2.4) 是由第 12 章中的扰动展开式识别的.

接下来运用 (2.4.6) 定义的空间 H, 并用它定义空间

$$X = H_0^1(\Omega) \times L^2(\Omega; H).$$

这是一个 Hilbert 空间且有内积

$$(U, V)_X = (\nabla u, \nabla v)_{L^2(\Omega)} + (\nabla_y u_1, \nabla_y v_1)_{L^2(\Omega\times\mathbb{T}^d)}.$$

对所有的 $U=\{u,u_1\}$, $V=\{v,v_1\}$, 并推导范数

$$\|U\|_X^2=\|\nabla u\|_{L^2(\Omega)}^2+\|\nabla_y u_1\|_{L^2(\Omega\times\mathbb{T}^d)}^2.$$

这被证明是正确的空间, 可以在它上面对双尺度极限进行描述, 接下来进行说明.

引理 19.4　设 $u^\varepsilon(x)$ 是满足定理 19.1 假设条件下的 (19.2.1) 的解, 则存在函数 $\{u(x),u_1(x,y)\}\in X$, 使得沿着一个子集, u^ε 和 ∇u^ε 分别双尺度收敛于 $u(x)$ 和 $\nabla_x u+\nabla_y u_1$.

证明　由定理 7.5 可知 $\|u^\varepsilon\|_{H_0^1(\Omega)}\leqslant C$, 这暗含了在 $H_0^1(\Omega)$ 上存在一个子序列弱收敛于极限 u. 由定理 2.36 可知存在函数 $u\in H_0^1(\Omega)$, $u_1\in L^2(\Omega;H)$, 使得在一个更小的子序列上可能有

$$u^\varepsilon\xrightarrow{2}u,\tag{19.3.2a}$$

$$\nabla u^\varepsilon\xrightarrow{2}\nabla_x u+\nabla_y u_1.\tag{19.3.2b}$$

接下来说明 $\{u,u_1\}$ 满足双尺度系统

$$-\nabla_y\cdot(A(y)(\nabla_x u+\nabla_y u_1))=0,\ \text{在}\Omega\times\mathbb{T}^d\text{上},\tag{19.3.3a}$$

$$-\nabla_x\cdot\left(\int_{\mathbb{T}^d}A(y)(\nabla_x u+\nabla_y u_1)\mathrm{d}y\right)=f,\ \text{在}\Omega\text{上},\tag{19.3.3b}$$

$$u(x)=0,\quad x\in\partial\Omega,\ u_1(x,y)\text{在}y\text{上具周期性}.\tag{19.3.3c}$$

通过定义双尺度系统的弱形式可以说明双线性形式

$$a[U,\Phi]=\int_\Omega\int_{\mathbb{T}^d}\langle A(\nabla_x u+\nabla_y u_1),\nabla_x\phi_0+\nabla_y\phi_1\rangle\mathrm{d}y\mathrm{d}x,$$

其中 $\Phi:=\{\phi_0,\phi_1\}\in X$. 双尺度系统 (19.3.3) 的弱形式是为了寻求 $U\in X$, 使得

$$a[U,\Phi]=(f,\phi_0),\quad\forall\Phi\in X.\tag{19.3.4}$$

为了明白这个, 设 $\phi_0=0$ 可得

$$\int_\Omega\int_{\mathbb{T}^d}\langle A(\nabla_x u+\nabla_y u_1),\nabla_y\phi_1\rangle\mathrm{d}y\mathrm{d}x=0.$$

(19.3.3a) 的弱形式是很精确的. 接下来, 在 (19.3.4) 中令 $\phi_1=0$ 可得

$$\int_\Omega\int_{\mathbb{T}^d}\langle A(\nabla_x u+\nabla_y u_1),\nabla_x\phi_0\rangle\mathrm{d}y\mathrm{d}x=(f,\phi_0),$$

这是 (19.3.3b) 的弱形式. 边界条件 (19.3.3c) 遵从这样一个事实, 即 $u\in H_0^1(\Omega)$ 和 $u_1\in L^2(\Omega;H)$.

引理 19.5 设 $u^\varepsilon(x)$ 是 (19.2.1) 的弱解并满足定理 19.1 的假设, 则引理 19.4 中的任意极限点 $\{u, u_1\}$ 是双尺度系统 (19.3.3) 的弱解.

证明 (19.2.1) 的弱形式是为了找到 $u^\varepsilon \in H_0^1(\Omega)$, 使得

$$\int_\Omega \langle A^\varepsilon \nabla u^\varepsilon, \nabla \phi^\varepsilon \rangle \mathrm{d}x = (f, \phi^\varepsilon), \quad \forall \phi^\varepsilon \in H_0^1(\Omega). \tag{19.3.5}$$

对 $\phi_0 \in C_0^\infty(\Omega)$, $\phi_1 \in C_0^\infty(\Omega; C_{\mathrm{per}}^\infty(\mathbb{T}^d))$, 运用 (19.3.1) 形式的测度函数, 可清晰地得到 $\phi^\varepsilon \in H_0^1(\Omega)$. 使用 (19.3.5) 中的测度函数, 并重新组合各项可得

$$I_1 + \varepsilon I_2 = (f, \phi_0 + \varepsilon \phi_1),$$

其中

$$I_1 = \int_\Omega \left\langle \nabla u^\varepsilon, (A^\varepsilon)^{\mathrm{T}} \left(\nabla_x \phi_0(x) + \nabla_y \phi_1 \left(x, \frac{x}{\varepsilon} \right) \right) \right\rangle \mathrm{d}x,$$

$$I_2 = \int_\Omega \left\langle \nabla u^\varepsilon, (A^\varepsilon)^{\mathrm{T}} \nabla_x \phi_1 \left(x, \frac{x}{\varepsilon} \right) \right\rangle \mathrm{d}x.$$

函数$A^\varepsilon(x)^{\mathrm{T}} \left(\nabla_x \phi_0(x) + \nabla_y \phi_1 \left(x, \frac{x}{\varepsilon} \right) \right)$ 是 $\psi(x,y) := \psi_1(y)\psi_2(x,y)(y = x/\varepsilon)$ 的形式, 其中 $\psi_1 \in L^\infty(\mathbb{T}^d)$ 和 $\psi_2 \in L^2(\Omega; C_{\mathrm{per}}(\mathbb{T}^d))$. 因此, 由引理 2.31, 可以将 $\psi(x,y)$ 作为一个测度函数, 并通过双尺度极限可得

$$I_1 \to \int_\Omega \int_{\mathbb{T}^d} \langle A(y)(\nabla_x u + \nabla_y u_1), (\nabla_x \phi_0 + \nabla_y \phi_1) \rangle \mathrm{d}y \mathrm{d}x, \quad \varepsilon \to 0.$$

函数 $A^\varepsilon(x)^{\mathrm{T}} \nabla_x \phi_1(x,y)$ 也是一个容许测度函数. 类似地, 在 I_2 中通过双尺度极限可得 $\varepsilon I_2 \to 0$. 此外, 由定理 2.29 可知在 $L^2(\Omega)$ 中 $\phi_0 + \varepsilon \phi_1 \rightharpoonup \phi_0$ 是弱收敛, 这暗含了

$$(f, \phi_0 + \varepsilon \phi_1) \to (f, \phi_0).$$

总结上述可得极限方程

$$\int_\Omega \int_{\mathbb{T}^d} \langle A(y)(\nabla_x u + \nabla_y u_1), \nabla_x \phi_0 + \nabla_y \phi_1 \rangle \mathrm{d}y \mathrm{d}x = (f, \phi_0),$$

从而可推导 (19.3.4). 在推导这个等式的过程中, 假设 ϕ_0, ϕ_1 是光滑的, 一个密度参数可以使得对于每一个 $\phi_0 \in H_0^1(\Omega)$, $\phi_1 \in L^2(\Omega; H)$, (19.3.4) 都成立. □

引理 19.6 在定理 19.1 的假设下, 双尺度系统 (19.3.3) 有唯一弱解 $\{u, u_1\} \in X$.

证明 将使用 Lax-Milgram 定理. 双尺度系统的弱形式由 (19.3.4) 给出, 不得不验证双线性形式 a 是连续且椭圆性的. 为了建立这两种性质. 设 $a = a(x)$ 和 $b = b(x,y)$ 是光滑函数, b 在 y 上具有周期性, 则

$$\int_\Omega \int_{\mathbb{T}^d} |\nabla_x a + \nabla_y b|^2 \mathrm{d}y \mathrm{d}x = \|\{a, b\}\|_X^2.$$

上述等式遵从如下事实:

$$\int_\Omega \int_{\mathbb{T}^d} \langle \nabla_x a, \nabla_y b \rangle \mathrm{d}y\mathrm{d}x = \int_\Omega \left(\int_{\mathbb{T}^d} \nabla_y \cdot (b\nabla_x a)\mathrm{d}y \right) \mathrm{d}x = 0,$$

运用散度定理和 $b\nabla_x a$ 在 y 上的周期性 (见注 7.13) 可得.

以连续性开始, 运用 L^∞ 在 $A(y)$ 上有界以及 Cauchy-Schwarz 不等式可得

$$\begin{aligned} a[U, \Phi] =& \int_\Omega \int_{\mathbb{T}^d} \langle A(\nabla_x u + \nabla_y u_1), \nabla_x \phi_0 + \nabla_y \phi_1 \rangle \mathrm{d}y\mathrm{d}x \\ \leqslant& \beta \int_\Omega \int_{\mathbb{T}^d} |\nabla_x u + \nabla_y u_1| |\nabla_x \phi_0 + \nabla_y \phi_1| \mathrm{d}y\mathrm{d}x \\ \leqslant& \beta \|U\|_X \|\Phi\|_X. \end{aligned}$$

接着进行椭圆性的证明. 使用散度定理和周期性可得

$$\begin{aligned} a[U, U] =& \int_\Omega \int_{\mathbb{T}^d} \langle A(\nabla_x u + \nabla_y u_1), \nabla_x u + \nabla_y u_1 \rangle \mathrm{d}y\mathrm{d}x \\ \geqslant& \alpha \int_\Omega \int_{\mathbb{T}^d} |\nabla_x u + \nabla_y u_1|^2 \mathrm{d}y\mathrm{d}x \\ =& \alpha \|U\|_X^2. \end{aligned}$$

因此,

$$a[U, U] \geqslant \alpha \|U\|_X^2.$$

于是双线性形式 $a[U, \Phi]$ 是连续的且满足椭圆性条件, 并根据 Lax-Milgram 定理可证明 X 中双尺度系统的解存在且唯一.

接下来叙述形如第 12 章中均匀化方程形式的双尺度系统.

引理 19.7 考虑双尺度系统 (19.3.3) 的唯一解 $\{u, u_1\} \in X$, 则 u 是均匀化方程 (19.2.2) 的唯一解, 并且 $u_1(x, y)$ 由下式给出:

$$u_1(x, y) = \chi(y) \cdot \nabla u(x), \tag{19.3.6}$$

其中 $\chi(y)$ 是单胞问题 (19.2.4) 的解.

证明 将 (19.3.6) 代入 (19.3.3a) 可得

$$-\nabla_y \cdot (\nabla_y \chi A^{\mathrm{T}}) \cdot \nabla_x u = (\nabla_y \cdot A^{\mathrm{T}}) \cdot \nabla_x u.$$

如果 $\chi \in H$ 是单胞问题 (19.2.4) 的解, 则满足这个方程. 方程 (19.3.3b) 变成

$$-\nabla_x \cdot \left(\int_{\mathbb{T}^d} A(\nabla_x u + (\nabla_y \chi)^{\mathrm{T}} \nabla_x u)\mathrm{d}y \right) = f,$$

使得

$$-\nabla_x \cdot \left(\int_{\mathbb{T}^d} A(I+\nabla_y \chi^{\mathrm{T}})\mathrm{d}y\right)\nabla_x u = f,$$

因此,

$$-\nabla_x \cdot (\bar{A}\nabla_x u) = f.$$

这是很精确的, 即具有均匀系数的均匀化方程由 (19.2.3) 给出.

只要 u_0 满足均匀化方程, 选取 (19.3.6) 得到的 u_1 就可以解决双尺度系统, 这一事实暗含了解决双尺度系统, 函数集 $\{u, u_1\}$ 是唯一可能的集合, 这是由已经证明了的存在性和唯一性所决定的. □

接下来归纳定理 19.1 的证明. 第一个引理显示了沿着一个子序列, u^ε 双尺度收敛于 $\{u, u_1\}$, 其中 u 独立于 y. 此外, u^ε 在 $H_0^1(\Omega)$ 上弱收敛, 而在 $L^2(\Omega)$ 上强收敛, 但是极限是唯一的, 正如将 Lax-Milgram 定理应用到双尺度系统的证明一样. 因此, 由子序列原理知整个序列必须收敛, 到此证明完成. □

19.4 证明：H^1 上的强收敛

定理 19.1 暗含了 u^ε 在 $L(\Omega)$ 上强收敛于 $u(x)$, 从而为了证明定理 19.2, 只需证明

$$\lim_{\varepsilon\to 0}\left\|\nabla u^\varepsilon(x) - \nabla\left(u(x)+\varepsilon u_1\left(x,\frac{x}{\varepsilon}\right)\right)\right\|_{L^2(\Omega,\mathbb{R}^d)} = 0,$$

或等效地证明

$$\lim_{\varepsilon\to 0}\left\|\nabla u^\varepsilon(x) - \left(\nabla u(x)+\varepsilon\nabla_x u_1\left(x,\frac{x}{\varepsilon}\right)+\nabla_y u_1\left(x,\frac{x}{\varepsilon}\right)\right)\right\|_{L^2(\Omega,\mathbb{R}^d)} = 0$$

就足够了. 由 $f\in L^2(\Omega)$ 可得 $u(x)\in H^2(\Omega)\cap H_0^1(\Omega)$. 因此, 由于通过假设 $\chi\in C^1_{\mathrm{per}}(\mathbb{T}^d)$ 可知 $u_1(x,y)=\chi(y)\cdot\nabla u(x)\in H^1(\Omega;C^1_{\mathrm{per}}(\mathbb{T}^d))$, 由定理 2.28 可知这暗含了

$$\left\|\nabla_x u_1\left(x,\frac{x}{\varepsilon}\right)\right\|_{L^2(\Omega;\mathbb{R}^d)} \leqslant C,$$

所以

$$\left\|\varepsilon\nabla_x u_1\left(x,\frac{x}{\varepsilon}\right)\right\|_{L^2(\Omega;\mathbb{R}^d)} \to 0.$$

因此, 只要证明如下等式就足够了：

$$\lim_{\varepsilon\to 0}\left\|\nabla u^\varepsilon(x) - \left(\nabla u(x)+\nabla_y u_1\left(x,\frac{x}{\varepsilon}\right)\right)\right\|_{L^2(\Omega;\mathbb{R}^d)} = 0.$$

A 的一致椭圆型暗含了

$$
\begin{aligned}
&\alpha\left\|\nabla u^{\varepsilon}(x)-\left(\nabla u(x)+\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right)\right\|_{L^{2}(\Omega;\mathbb{R}^{d})}^{2}\\
=&\alpha\int_{\Omega}\left|\nabla u^{\varepsilon}(x)-\left(\nabla u(x)+\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right)\right|^{2}\mathrm{d}x\\
\leqslant&\int_{\Omega}\left\langle A\left(\frac{x}{\varepsilon}\right)\left(\nabla_{x}u^{\varepsilon}(x)-\nabla_{x}u(x)-\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right),\right.\\
&\left.\nabla_{x}u^{\varepsilon}(x)-\nabla_{x}u(x)-\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right\rangle\mathrm{d}x\\
\leqslant&\int_{\Omega}\left\langle A\left(\frac{x}{\varepsilon}\right)\nabla_{x}u^{\varepsilon}(x),\nabla_{x}u^{\varepsilon}(x)\right\rangle\mathrm{d}x\\
&+\int_{\Omega}\left\langle A\left(\frac{x}{\varepsilon}\right)\left(\nabla_{x}u(x)+\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right),\nabla_{x}u(x)+\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right\rangle\mathrm{d}x\\
&-\int_{\Omega}\langle\nabla_{x}u^{\varepsilon},(A^{\varepsilon}+(A^{\varepsilon})^{\mathrm{T}})(\nabla_{x}u+\nabla_{y}u_{1})\rangle\mathrm{d}x\\
=:&(f,u^{\varepsilon})+I_{1}^{\varepsilon}+I_{2}^{\varepsilon}.
\end{aligned}
$$

由定理 19.1 知 u^{ε} 在 $L^{2}(\Omega)$ 上强收敛于 u. 由强收敛暗含了弱收敛可得

$$(f,u^{\varepsilon})\to(f,u)=a[U,U],$$

其中最后一个等式遵从 (19.3.4). 此外, 由于

$$A\in L_{\mathrm{per}}^{\infty}(\mathbb{T}^{d};\mathbb{R}^{d\times d}),\quad u\in H^{2}(\Omega)\cap H_{0}^{1}(\Omega)$$

和 $\chi\in C_{\mathrm{per}}^{1}(\mathbb{T}^{d})$, 引理 2.31 暗含了

$$A\left(\frac{x}{\varepsilon}\right)\left(\nabla_{x}u(x)+\nabla_{y}u_{1}\left(x,\frac{x}{\varepsilon}\right)\right)$$

和

$$(A^{\varepsilon}+(A^{\varepsilon})^{\mathrm{T}})(\nabla_{x}u+\nabla_{y}u_{1}),$$

可以分别被用作 I_{1}^{ε} 和 I_{2}^{ε} 上的测度函数. 通过这两个表达式中的双尺度极限可得当 $\varepsilon\to0$ 时,

$$I_{1}^{\varepsilon}\to\int_{\Omega}\int_{\mathbb{T}^{d}}\langle A(y)(\nabla_{x}u(x)+\nabla_{y}u_{1}(x,y)),\nabla_{x}u(x)+\nabla_{y}u_{1}(x,y)\rangle\mathrm{d}y\mathrm{d}x$$

和

$$I_{2}^{\varepsilon}\to-\int_{\Omega}\int_{\mathbb{T}^{d}}\langle(A(y)+(A(y))^{\mathrm{T}})(\nabla_{x}u(x)+\nabla_{y}u_{1}(x,y)),\nabla_{x}u(x)+\nabla_{y}u_{1}(x,y)\rangle\mathrm{d}y\mathrm{d}x,$$

从而证明了

$$\lim_{\varepsilon\to 0}\alpha\left\|\nabla u^\varepsilon(x)-\left(\nabla u(x)+\nabla_y u_1\left(x,\frac{x}{\varepsilon}\right)\right)\right\|^2_{L^2(\Omega;\mathbb{R}^d)}\leqslant \alpha[U,U]+\alpha[U,U]-2\alpha[U,U]=0.$$

因此,

$$\lim_{\varepsilon\to 0}\left\|\nabla u^\varepsilon(x)-\left(\nabla u(x)+\nabla_y u_1\left(x,\frac{x}{\varepsilon}\right)\right)\right\|_{L^2(\Omega;\mathbb{R}^d)}=0,$$

定理证明完毕. □

19.5 讨论和参考

定理 19.1 和定理 19.2 的证明源于文献 [6]. 校正结论中的正则性假设, 定理 19.2, 都并非是最理想的, 见 [6, 定理 2.6] 或者 [66,9.12]. 然而, 需要某些正则性才能证明校正结论.

均匀化定理的证明主要包括两个步骤: 第 1 步是证明一个极限的存在性, 则需尽力使这个极限特征化. 一个极限的存在性遵从先验估计, 这些先验估计非常普通也并不难得到. 另一方面, 如极限的特征, 严格的合理性, 即用序列 u^ε 的极限满足均匀化偏微分方程 —— 通常是一个更为精细的问题. 使用适当的测度函数来特征化函数的一个序列的极限, 这一思想在偏微分方程理论中很常见, 见文献 [95] 以及相关参考文献. 在均匀化的内容中, 式 (19.3.1) 的测度函数已经在文献 [181] 中被 Kurtz 运用, 也可见 Evans 扰动测度函数方法[96,97].

Tartar 振荡测度函数法是基于使用单胞问题构造适当的测试函数, 见文献 [66, ch.8, 307; 311] 以及相应的参考. 椭圆型偏微分方程的均匀化结论也可以通过使用变量计算法和 Γ 收敛的概念来证明, 见文献 [214]. 这些方法适用于非线性偏微分方程的各种均匀化问题, 使用双尺度收敛的方法直接证明单调算子的一个均匀化定理是相对简洁的, 见练习第 7 题.

第 16 章 Markov 链平均化定理的证明和第 20 章抛物型偏微分方程均匀化定理的证明中用到的步步为营的技巧也可以用来证明椭圆型偏微分方程均匀化定理, 见本章练习的第 6 题和第 20 章练习的第 4 题. 这个方法可以发现误差估计, 见文献 [66,ch.7].

解耦双尺度系统并获得均匀化方程中双尺度展开式第一项的一个近似方程并不总是可能的. 为了证明这个, 需要一个关于非均匀化偏微分方程的 H^1 估计. 这可以归纳出双尺度极限是独立于微尺度的, 因此, 均匀化方程实际上是存在的. 这样一种状况对多尺度问题并不是总会发生的, 如见第 21 章线性运输偏微分方程的研究. 其他的例子可以在文献 [5, 6] 中找到. 从这一观点出发, 双尺度系统比均匀化方程更具基础性, 这是由于前者是方程的适定系统, 而后者甚至可能都不存在.

19.6 练　　习

1. 设 $a : \mathbb{T}^d \mapsto \mathbb{R}^d$ 是一个光滑、1 周期、平均值为零的无散场, 并且考虑稳态对流扩散方程的均匀化问题

$$-\Delta u^\varepsilon + \frac{1}{\varepsilon} a\left(\frac{x}{\varepsilon}\right) \cdot \nabla u^\varepsilon = f, \quad x \in \Omega, \tag{19.6.1a}$$

$$u^\varepsilon(x) = 0, \quad x \in \partial\Omega. \tag{19.6.1b}$$

运用双尺度收敛法证明这个偏微分方程的均匀化定理.

2. 对偏微分方程

$$-\Delta u^\varepsilon + \frac{1}{\varepsilon} a\left(\frac{x}{\varepsilon}\right) \cdot \nabla u^\varepsilon + V\left(\frac{x}{\varepsilon}\right) u^\varepsilon = f, \quad x \in \Omega,$$

$$u^\varepsilon(x) = 0, \quad x \in \partial\Omega$$

进行如第 1 题中同样的程序, 其中 $V(y)$ 是光滑和 1 周期且 $\int_{y\in\mathbb{T}^d} V(y)\mathrm{d}y > 0$.

3. 对偏微分方程

$$-\Delta u^\varepsilon + \frac{1}{\varepsilon} a\left(\frac{x}{\varepsilon}\right) \cdot \nabla u^\varepsilon + \frac{1}{\varepsilon} V\left(\frac{x}{\varepsilon}\right) u^\varepsilon = f, \quad x \in \Omega,$$

$$u^\varepsilon(x) = 0, \quad x \in \partial\Omega$$

进行如第 2 题中同样的程序, 其中 $V(y)$ 是光滑和 1 周期的且均值为零.

4. 对偏微分方程 (19.6.1) 进行如第 3 题中同样的程序, 并且

$$a(y) = -\nabla_y U(y),$$

其中 U 是光滑和 1 周期的.

5. 陈述并证明 Neumann 问题的一个均匀化定理, 这个问题是通过交换 (19.2.1) 中 Dirichlet 边界条件和均匀化 Neumann 边界条件发现的. 运用双尺度收敛证明这个结论.

6. 运用适当的能量估计证明椭圆偏微分方程的均匀化定理. (见第 20 章, 这个定理是在偏微分方程中进行的.)

7. (单调算子均匀化) 考虑如下边界值问题:

$$-\nabla \cdot \left(a\left(\frac{x}{\varepsilon}, \nabla u^\varepsilon\right)\right) = f, \quad x \in \Omega, \tag{19.6.2a}$$

$$u^\varepsilon = 0, \quad x \in \partial\Omega. \tag{19.6.2b}$$

假定 $f \in L^2(\Omega)$, 并且对函数 $a(y,\lambda) : \mathbb{T}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ 作如下假设:

(i) 映射 $\lambda \to a(y,\lambda)$ 是可测的且在 y 上对每个 λ 都是 1 周期的;

(ii) 映射 $y \to a(y,\lambda)$ 在 $y \in Y$ 上是连续的;

(iii) 存在一个 $c > 0$, 使得

$$c|\lambda|^2 \leqslant a(y,\lambda) \cdot \lambda, \quad \forall y \in Y, \forall \lambda \in \mathbb{R}^d; \tag{19.6.3}$$

(iv) 存在一个 $c > 0$, 使得

$$|a(y,\lambda)| \leqslant c(1+|\lambda|), \quad \forall y \in Y, \forall \lambda \in \mathbb{R}^d; \tag{19.6.4}$$

(v) $a(y,\lambda)$ 是强单调的;

$$[a(y,\lambda) - a(y,\mu)] \geqslant c|\lambda - \mu|^2, \quad \forall y \in Y, \quad \forall \lambda, \mu \in \mathbb{R}^d. \tag{19.6.5}$$

(a) 陈述并证明边界值问题 (19.6.2) 的解存在性和唯一性定理 (如果需要可参见文献 [339, 2.14 节]);

(b) 运用双尺度收敛陈述并证明 (19.6.2) 的一个均匀化定理;

(c) 陈述并证明一个校正型结论.

第 20 章　抛物型偏微分方程的均匀化：收敛性定理

20.1　引　　言

本章证明了第 13 章研究的二阶抛物型偏微分方程形式的均匀化定理. 证明的方法在构造上与第 16 章中用到的非常类似, 即对 Markov 链的一个平均化定理的证明. 精确地说, 在多尺度展开式中得到一个误差方程, 并可以直接从这个方程中估计误差. 第 16 章中用到的重要估计遵从这样一个事实, 即本章中的 Q 是随机矩阵, 本章中的类似估计遵从抛物型偏微分方程的极大值原理; 20.2 节陈述了基本结论; 20.3 节是证明; 本章提出的均匀化定理的拓展和参考文献在 20.4 节.

20.2　定　　理

如第 13 章, 设 $b^\varepsilon = b(x/\varepsilon)$. 考虑方程 (13.2.4), 即

$$\frac{\partial u^\varepsilon}{\partial t} = \frac{1}{\varepsilon} b^\varepsilon \cdot \nabla u^\varepsilon + D\Delta u^\varepsilon, \quad (x,t) \in \mathbb{R}^d \times (0,T),$$

$$u^\varepsilon = f, \quad (x,t) \in \mathbb{R}^d \times \{0\}.$$

回顾生成子

$$\mathcal{L}_0 = b(y) \cdot \nabla_x + D\Delta_y,$$

具有周期边界条件且它的 L^2 伴随 $\mathcal{L}_0^*$ 也具有周期边界条件, 不变分布是在 $\mathcal{L}_0^*$ 的零空间上, 有效扩散率由 (13.3.2) 给出

$$\mathcal{K} = DI + 2D\int_{\mathbb{T}^d} \nabla_y \chi(y)^{\mathrm{T}} \rho(y)\mathrm{d}y + \int_{\mathbb{T}^d} (b(y) \otimes \chi(y))\rho(y)\mathrm{d}y,$$

其中 ρ 是 $\mathcal{L}_0^*$ 上零空间的不变分布, 由 (13.2.6) 给出. 回顾 χ 是如下单胞问题的解:

$$-\mathcal{L}_0\chi(y) = b(y), \quad \chi\text{是 1 周期}, \int_{\mathcal{Y}} \chi(y)\rho(y)\mathrm{d}y = 0. \tag{20.2.1}$$

定理 20.1　设 $u^\varepsilon(x,t)$ 是 (13.2.4) 的解, 其中 $b \in C^\infty_{\mathrm{per}}(\mathbb{T}^d)$ 且 $f \in C^\infty_b(\mathbb{R}^d)$. 设 $u(x,t)$ 是如下均匀化方程的解:

$$\frac{\partial u}{\partial t} = \mathcal{K} : \nabla\nabla u, \quad (x,t) \in \mathbb{R}^d \times (0,T), \tag{20.2.2a}$$

$$u = f(x), \quad (x,t) \in \mathbb{R}^d \times \{0\}, \tag{20.2.2b}$$

则

$$\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}^d \times (0,T))} \leqslant C\varepsilon, \tag{20.2.3}$$

从而在 $L^\infty(\mathbb{R}^d \times (0,T))$ 上有 $u^\varepsilon \to u$.

20.3 证　明

第 13 章推导了双尺度展开式

$$u^\varepsilon(x,t) \approx u(x,t) + \varepsilon u_1^\varepsilon(x,t) + \varepsilon^2 u_2^\varepsilon(x,t),$$

其中 $u_i^\varepsilon(x,t) := u_i(x, x/\varepsilon, t)(i=1,2)$, 并且

$$u_1(x,y,t) = \chi(y) \cdot \nabla_x u(x,t). \tag{20.3.1}$$

一个与第 12 章提出的类似分析可以得到 u_2 的表达式 (见第 13 章练习第 1 题), 即

$$u_2(x,y,t) = \Theta(y) : \nabla_x \nabla_x u(x,t). \tag{20.3.2}$$

向量场 $\chi(y)$ 是向量值 Poisson 方程 (20.2.1) 的解, 并且矩阵场 $\Theta(y)$ 是如下矩阵值 Poisson 方程的解:

$$\begin{aligned} -\mathcal{L}_0 \Theta = {} & b(y) \otimes \chi(y) + 2D\nabla_y \chi(y)^{\mathrm{T}} \\ & - \int_{\mathbb{T}^d} (b(y) \otimes \chi(y) + 2D\nabla_y \chi(y)^{\mathrm{T}}) \rho(y) \mathrm{d}y, \end{aligned} \tag{20.3.3}$$

并且具有周期边界条件. 校正场 $\chi(y)$, $\Theta(y)$ 是具有光滑系数和周期边界条件的一致椭圆型偏微分方程的解, 因此, 由椭圆正则性定理, χ 和 Θ 都具有自身的各阶导数, 并且对于每一个整数 $k \geqslant 0$ 有界,

$$\|\chi\|_{C^k(\mathbb{R}^d;\mathbb{R}^d)} \leqslant C, \quad \|\Theta\|_{C^k(\mathbb{R}^d;\mathbb{R}^{d\times d})} \leqslant C. \tag{20.3.4}$$

此外, 对初始条件 f 的假设暗含了 u 作为一个具有常数系数的偏微分方程的解, 属于 $L^\infty(\mathbb{R}^d \times (0,T))$, 并且它所有的导数都与时间和空间有关. 这个事实并结合估计 (20.3.4) 提供了下面这样一个边界:

$$\|u_1^\varepsilon\|_{L^\infty((0,T)\times\mathbb{R}^d)} \leqslant C, \quad \|u_2^\varepsilon\|_{L^\infty((0,T)\times\mathbb{R}^d)} \leqslant C, \tag{20.3.5}$$

其中常数 C 独立于 ε. 在前面, 使用符号 $u_i^\varepsilon(x,t) := u_i(x, x/\varepsilon, t)(i=1,2)$.

设 $\mathcal{R}^\varepsilon(x,t)$ 表示由方程

$$u^\varepsilon(x,t) = u(x,t) + \varepsilon u_1^\varepsilon(x,t) + \varepsilon^2 u_2^\varepsilon(x,t) + \mathcal{R}^\varepsilon(x,t) \tag{20.3.6}$$

定义的余项, 定义

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon} b^\varepsilon \cdot \nabla_x + D\Delta_x.$$

将 $\mathcal{L}^\varepsilon$ 应用到形式为 $f(x, x/\varepsilon, t)$ 的函数中可得

$$\begin{aligned}\mathcal{L}^\varepsilon =& \frac{1}{\varepsilon^2}(b(y)\cdot\nabla_y + D\Delta_y) + \frac{1}{\varepsilon}(b(y)\cdot\nabla_x + 2D\nabla_x\nabla_y) + D\Delta_x \\ =& \frac{1}{\varepsilon^2}\mathcal{L}_0 + \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_2,\end{aligned}$$

其中 $y = x/\varepsilon$, 并且

$$\mathcal{L}_0 = b(y)\cdot\nabla_y + D\Delta_y,$$

$$\mathcal{L}_1 = b(y)\cdot\nabla_x + 2D\nabla_x\cdot\nabla_y,$$

$$\mathcal{L}_2 = D\Delta_x.$$

回顾 u_0, u_1 和 u_2 的构造, 使得

$$O\left(\frac{1}{\varepsilon^2}\right): \quad \mathcal{L}_0 u_0 = 0;$$

$$O\left(\frac{1}{\varepsilon}\right): \quad \mathcal{L}_0 u_1 = -\mathcal{L}_1 u_0;$$

$$O(1): \quad \mathcal{L}_0 u_2 = -\mathcal{L}_1 u_1 - \mathcal{L}_2 u_0 + \frac{\partial u_0}{\partial t}.$$

由于 $u_0 = u(x,t)$, 将 $\mathcal{L}^\varepsilon$ 应用到展开式 (20.3.6) 可得

$$\begin{aligned}\mathcal{L}^\varepsilon u^\varepsilon =& \mathcal{L}^\varepsilon(u + \varepsilon u_1 + \varepsilon^2 u_2) + \mathcal{L}^\varepsilon\mathcal{R}^\varepsilon \\ =& \frac{1}{\varepsilon^2}\mathcal{L}_0 u + \frac{1}{\varepsilon}(\mathcal{L}_0 u_1 + \mathcal{L}_1 u) + (\mathcal{L}_0 u_2 + \mathcal{L}_1 u_1 + \mathcal{L}_2 u) \\ &+ \varepsilon(\mathcal{L}_1 u_2 + \mathcal{L}_2 u_1) + \varepsilon^2\mathcal{L}_2 u_2 + \mathcal{L}^\varepsilon\mathcal{R}^\varepsilon \\ =& \frac{\partial u}{\partial t} + \varepsilon(\mathcal{L}_1 u_2 + \mathcal{L}_2 u_1) + \varepsilon^2\mathcal{L}_2 u_2 + \mathcal{L}^\varepsilon\mathcal{R}^\varepsilon.\end{aligned}$$

另一方面,

$$\frac{\partial u^\varepsilon}{\partial t} = \frac{\partial u}{\partial t} + \varepsilon\frac{\partial u_1}{\partial t} + \varepsilon^2\frac{\partial u_2}{\partial t} + \frac{\partial \mathcal{R}^\varepsilon}{\partial t}.$$

结合这两个方程和非均匀化方程可得

$$\frac{\partial \mathcal{R}^\varepsilon}{\partial t} = \mathcal{L}^\varepsilon\mathcal{R}^\varepsilon + \varepsilon F^\varepsilon(x,t),$$

其中

$$F^\varepsilon(x,t)=F(x,x/\varepsilon,t), \tag{20.3.7}$$

$$F(x,y,t)=\mathcal{L}_1u_2+\mathcal{L}_2u_1-\frac{\partial u_1}{\partial t}+\varepsilon\left(\mathcal{L}_2u_2-\frac{\partial u_2}{\partial t}\right). \tag{20.3.8}$$

此外,

$$\begin{aligned}f(x)=&u^\varepsilon(x,0)\\=&u(x,0)+\varepsilon u_1\left(x,\frac{x}{\varepsilon},0\right)+\varepsilon^2u_2\left(x,\frac{x}{\varepsilon},0\right)+\mathcal{R}^\varepsilon(x,0).\end{aligned}$$

因此, 计算 (20.2.2b) 可得

$$\mathcal{R}^\varepsilon(x,0)=\varepsilon h^\varepsilon(x),$$

其中

$$h^\varepsilon(x,0)=-u_1\left(x,\frac{x}{\varepsilon},0\right)-\varepsilon u_2\left(x,\frac{x}{\varepsilon},0\right). \tag{20.3.9}$$

将这些计算放在一起可得余项 $\mathcal{R}^\varepsilon(x,t)$ 的如下 Cauchy 问题:

$$\frac{\partial\mathcal{R}^\varepsilon}{\partial t}=\mathcal{L}^\varepsilon\mathcal{R}^\varepsilon+\varepsilon F^\varepsilon(x,t),\quad (x,t)\in\mathbb{R}^d\times(0,T), \tag{20.3.10a}$$

$$\mathcal{R}^\varepsilon=\varepsilon h^\varepsilon(x),\quad (x,t)\in\mathbb{R}^d\times\{0\}, \tag{20.3.10b}$$

其中 $F^\varepsilon(x,t)$ 和 $h^\varepsilon(x)$ 分别由 (20.3.7) 和 (20.3.9) 给出.

为了证明定理 20.1, 需要关于 F^ε 和 h^ε 的估计.

引理 20.2　在定理 20.1 假设下, $F^\varepsilon(x,t)$ 和 $h^\varepsilon(x,t)$ 分别满足

$$\|F^\varepsilon\|_{L^\infty(\mathbb{R}^d\times(0,T))}\leqslant C \tag{20.3.11}$$

和

$$\|h^\varepsilon\|_{L^\infty(\mathbb{R}^d)}\leqslant C, \tag{20.3.12}$$

其中常数 C 独立于 ε.

证明　已经知道

$$\begin{aligned}F(x,y,t)=&\mathcal{L}_1(\Theta(y):\nabla_x\nabla_xu(x,t))+\mathcal{L}_2(\chi(y)\cdot\nabla_xu(x,t))\\&-\frac{\partial}{\partial t}(\chi(y)\cdot\nabla_xu(x,t))+\varepsilon(D\Delta_x\Theta(y):\nabla_x\nabla_xu(x,t)\\&-\frac{\partial}{\partial t}(\Theta(y):\nabla_x\nabla_xu(x,t))\Big).\end{aligned} \tag{20.3.13}$$

估计 (20.3.4) 中 L^∞ 在 u 上有界并具有各阶导数, 这暗含了 (20.3.11).

此外,

$$h^\varepsilon(x)=\chi\left(\frac{x}{\varepsilon}\right)\cdot\nabla_xu(x,0)+\varepsilon\Theta\left(\frac{x}{\varepsilon}\right):\nabla_x\nabla_xu(x,0).$$

关于 $\chi(y)$ 和 $\Theta(y)$ 的一致估计, 以及对初始条件 (13.2.4) 的假设导致了估计 (20.3.12). □

定理 20.1 的证明 余项 $\mathcal{R}^\varepsilon(x,t)$ 满足方程 (20.3.10a). 使用第 7 章中的极大值原理估计 (7.3.6) 可得

$$\begin{aligned}\|\mathcal{R}^\varepsilon\|_{L^\infty(\mathbb{R}^d\times(0,T))} &\leqslant \varepsilon\|h^\varepsilon\|_{L^\infty(\mathbb{R}^d)} + \varepsilon\int_0^T \|F^\varepsilon(\cdot,s)\|_{L^\infty(\mathbb{R}^d)}\mathrm{d}s \\ &\leqslant \varepsilon C + \varepsilon CT \\ &\leqslant C\varepsilon. \end{aligned} \tag{20.3.14}$$

结合 (20.3.6) 和 (20.3.14) 以及 (20.3.5), 并使用三角不等式可得

$$\begin{aligned}\|u^\varepsilon - u\|_{L^\infty(\mathbb{R}\times(0,T))} &= \|\varepsilon u_1 + \varepsilon^2 u_2 + \mathcal{R}^\varepsilon\|_{L^\infty(\mathbb{R}^d\times(0,T))} \\ &\leqslant \varepsilon\|u_1\|_{L^\infty(\mathbb{R}^d\times(0,T))} + \varepsilon^2\|u_2\|_{L^\infty(\mathbb{R}^d\times(0,T))} \\ &\quad + \|\mathcal{R}^\varepsilon\|_{L^\infty(\mathbb{R}^d\times(0,T))} \leqslant C\varepsilon,\end{aligned}$$

从上式知 (20.2.3) 成立. □

20.4 讨论和参考

均匀化定理的证明依靠极大值原理, 也许这是最简单的方法, 只要对偏微分方程系数作充分的正定性假设, 在无界域偏微分方程将得到很好的研究. 然而, 其他方法也可以用来研究抛物型偏微分方程的均匀化. 这些包括概率法[248]、能量法[33,ch.2] 以及双尺度收敛法[5]. 运用黏度解理论, 基于极大值原理的估计可以用来证明非线性 (甚至不是完全非线性) 椭圆型偏微分方程和抛物型偏微分方程的均匀化定理, 如可见文献 [95~97, 203].

本章用到的方法, 即误差项的一个偏微分方程的导数和使用先验估计来控制推动误差方程的数据, 可以被应用到奇异扰动理论的各种问题中, 其他的一些例子可见文献 [240, 246]. 这个方法通常被称为步步为营法. 对于具有随机系数的非线性抛物型偏微分方程, 有兴趣的一个主题就是波前解的存在和噪声对它们的影响, 见这个领域的参考文献 [333].

20.5 练　　习

1. 考虑 $\nabla\cdot b(x) = 0$ 的情况. 设 Cauchy 问题 (13.2.4) 的解光滑有界且在无穷远处衰减得足够快.

(a) 考虑非均匀化 Cauchy 问题

$$\frac{\partial \mathcal{R}}{\partial t} = \frac{1}{\varepsilon} b^\varepsilon \cdot \nabla \mathcal{R} + D\Delta \mathcal{R} + F(x,t), \quad (x,t) \in \mathbb{R}^d \times (0,T),$$

$$\mathcal{R}(x,0) = f(x), \quad x \in \mathbb{R}^d,$$

其中 $f \in L^2(\mathbb{R}^d)$ 和 $F \in L^2((0,T)\times\mathbb{R}^d)$, 证明估计

$$\|\mathcal{R}\|^2_{L^2((0,T)\times\mathbb{R}^d)} + C_1\|\nabla\mathcal{R}\|^2_{L^2((0,T)\times\mathbb{R}^d)} \leqslant C_2\|f\|^2_{L^2(\mathbb{R}^d)} + C_3\|F\|^2_{L^2((0,T)\times\mathbb{R}^d)};$$

(b) 运用以上结论证明 $L^2((0,T)\times\mathbb{R}^d)$ 的收敛性;

(c) 均匀化定理成立的最大时间间隔 $(0,T)$ 是多少?

2. 对 Cauchy 问题

$$\frac{\partial u^\varepsilon}{\partial t} = \nabla \cdot \left(A\left(\frac{x}{\varepsilon}\right)\nabla u^\varepsilon\right), \quad (x,t) \in \mathbb{R}^d \times (0,T),$$

$$u^\varepsilon(x,0) = u_{\text{in}}(x), \quad x \in \mathbb{R}^d$$

进行与第 1 题同样的程序, 其中矩阵 $A(y)$ 满足标准周期性、光滑性和一致椭圆型假设.

3. 考虑初始边界值问题

$$\frac{\partial u^\varepsilon}{\partial t} = \nabla \cdot \left(A\left(\frac{x}{\varepsilon}\right)\nabla u^\varepsilon\right), \quad (x,t) \in \Omega \times (0,T),$$

$$u^\varepsilon(x,0) = u_{\text{in}}(x), \quad x \in \bar{\Omega},$$

$$u^\varepsilon(x,t) = 0, \quad (x,t) \in \partial\Omega \times [0,T],$$

其中 Ω 是 $\mathbb{R}^d$ 上的有界域并具有光滑边界, $A(y)$ 满足标准假设.

(a) 证明估计

$$\|u^\varepsilon\|^2_{L^\infty((0,T)\times\Omega)} + C\|u^\varepsilon\|^2_{L^2((0,T);H^1_0(\Omega))} \leqslant \|u_{\text{in}}\|^2_{L^2(\Omega)};$$

(b) 运用前一个估计, 用双尺度收敛的方法证明均匀化定理;

(c) 你能将步步为营法应用到这个问题吗?

4. 运用 7.2.4 小节中椭圆型偏微分方程的极大值原理和本章中的方法, 证明无散度形式下, 具有快速振荡系数的二阶非一致椭圆型偏微分方程的一个均匀化定理.

第 21 章　线性运输方程和抛物型偏微分方程的平均化：收敛性定理

21.1　引　　言

本章中证明了具有周期、无散度速度场的线性运输方程和线性抛物型偏微分方程的一个平均化定理. 分开研究抛物型情况 ($D > 0$) 和运输情况 ($D = 0$). 这是因为一般来说, 这两种情况下的平均化方程是不同的①.

21.2 节提出了两个收敛性定理; $D > 0$ 情况下平均化定理的证明在 21.3 节中; 21.4 节中说明了在 $D = 0$ 情况下, 当速度场在单位圆环上没有生成一个遍历性流时, 均匀化极限导致方程的一个双系统, 即双尺度系统, 两个定理的证明都是基于双尺度收敛方法; 在 21.5 节中以评论及参考备注结束本章.

21.2　定　　理

如第 14 章定义的:

$$b^{\varepsilon}(x) = b\left(\frac{x}{\varepsilon}\right).$$

由于分析局限于无散度场, 很自然地定义

$$\bar{b} = \int_{\mathbb{T}^d} b(y)\mathrm{d}y.$$

这是当不变密度 ρ 是一个常数函数时, 结论 14.1 给出的 $\bar{b}$ 的形式, 这种情况表示的是无散场.

回顾定义 7.24 和注 7.28 包含的运输方程弱解的定义, 这些定义也可以扩展到抛物型方程中. 本章的第一个定理就是有关抛物线型情况的.

定理 21.1　设 $u^{\varepsilon}(x,t)$ 是

$$\frac{\partial u^{\varepsilon}}{\partial t} - b^{\varepsilon} \cdot \nabla u^{\varepsilon} = \varepsilon D \Delta u^{\varepsilon}, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}^+, \tag{21.2.1a}$$

$$u^{\varepsilon} = g, \quad x \in \mathbb{R}^d \tag{21.2.1b}$$

---

① 在偏微分方程文献中, 本章建立的极限方程有时是作为均匀化方程被参考的, 而不是平均化方程. 然而, 如第 14 章讨论的, 选择使用在 1.3 节介绍的一致术语.

的弱解, 并假设 $D>0$, $g\in C_b^\infty(\mathbb{R}^d)$ 且 b 是光滑、无散度和 1 周期的, 则 u^ε 双尺度收敛于 $u_0\in L^2(\mathbb{R}^+\times\mathbb{R}^d)$, 其中 u_0 是

$$\frac{\partial u_0}{\partial t}-\bar{b}\cdot\nabla u_0=0,\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+, \tag{21.2.2a}$$

$$u_0=g,\quad (x,t)\in\mathbb{R}^d\times\{0\} \tag{21.2.2b}$$

的一个弱解. 接下来考虑 $D=0$ 的情况.

定理 21.2 设 $u^\varepsilon(x,t)$ 是

$$\frac{\partial u^\varepsilon}{\partial t}-b^\varepsilon\cdot\nabla u^\varepsilon=0,\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+, \tag{21.2.3a}$$

$$u^\varepsilon=g,\quad x\in\mathbb{R}^d \tag{21.2.3b}$$

的弱解, 并假设 $g\in C_b^\infty(\mathbb{R}^d)$ 且 b 是光滑、无散度和 1 周期的, 则 u^ε 双尺度收敛于 $u_0\in L^2(\mathbb{R}^+\times\mathbb{R}^d;L^2_{\mathrm{per}}(\mathbb{T}^d))$, 其中 u_0 满足双尺度系统

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\int_{\mathbb{T}^d} b(y)\cdot\nabla_y\phi(x,y,t)u_0(x,y,t)\mathrm{d}y\mathrm{d}x\mathrm{d}t=0, \tag{21.2.4a}$$

$$\phi\in C_0^\infty(\mathbb{R}\times\mathbb{R}^d;C^\infty_{\mathrm{per}}(\mathbb{T}^d)), \tag{21.2.4b}$$

$$\begin{aligned}&\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\int_{\mathbb{T}^d}\left(\frac{\partial\phi(x,y,t)}{\partial t}-b(y)\cdot\nabla_x\phi(x,y,t)\right)u_0(x,y,t)\mathrm{d}y\mathrm{d}x\mathrm{d}t\\&\quad+\int_{\mathbb{R}^d}g(x)\left(\int_{\mathbb{T}^d}\phi(x,y,0)\mathrm{d}y\right)\mathrm{d}x=0.\end{aligned} \tag{21.2.4c}$$

此外, 这些方程有唯一解.

注意到这个双尺度系统包括一个辅助函数 $\phi(x,y,t)\in C_0^\infty(\mathbb{R}\times\mathbb{R}^d;C^\infty_{\mathrm{per}}(\mathbb{T}^d))$. 由于是椭圆型情况, 为了得到多尺度展开式中第一项的一个近似方程, 需要解耦双尺度系统. 当函数 b 在单位圆环上生成一个遍历性流时, 解耦双尺度系统可得到一个平均化运输方程.

推论 21.3 设 b 是一个光滑、无散度向量场, 在 $\mathbb{T}^d$ 上生成一个遍历性流, 则双尺度极限独立于 y, 并且是如下 Cauchy 问题的一个弱解:

$$\frac{\partial u_0}{\partial t}-\bar{b}\cdot\nabla u_0=0,\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+, \tag{21.2.5a}$$

$$u_0=g,\quad (x,t)\in\mathbb{R}^d\times\{0\}. \tag{21.2.5b}$$

21.3 证明：$D>0$

注意到从定理 7.22, 由于 b^ε 和 f 都是光滑且有界的, Cauchy 问题 (21.1) 存在

唯一古典解. 此外, 当 $|x| \to \infty$ 时, 问题的解衰减到 0. 这个事实验证了如下分部积分:

首先获取一个能量估计. 在方程 (21.2.1a) 两边乘以 u^ε, 在 $\mathbb{R}^d$ 上积分, 运用 b^ε 是无散度的事实, 并且采用分部积分可得

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}|u^\varepsilon(\cdot,x)|^2\mathrm{d}x + D\varepsilon\int_{\mathbb{R}^d}|\nabla u^\varepsilon(\cdot,x)|^2\mathrm{d}x = 0.$$

接下来, 在 $(0,T)$ 上积分推导出

$$\int_{\mathbb{R}^d}|u^\varepsilon(\cdot,t)|^2\mathrm{d}x + 2D\varepsilon\int_0^T\int_{\mathbb{R}^d}|\nabla u^\varepsilon|^2\mathrm{d}x\mathrm{d}t = \int_{\mathbb{R}^d}|g|^2\mathrm{d}x.$$

因此,

$$\|u^\varepsilon\|^2_{L^\infty((0,T);L^2(\mathbb{R}^d))} + C\varepsilon\|\nabla u^\varepsilon\|^2_{L^2((0,T)\times\mathbb{R}^d)} \leqslant C. \tag{21.3.1}$$

注意到 T 在这个估计中是任意的.

定理 2.39 的思想, 部分 (iii), 估计 (21.3.1) 暗含了存在一个函数 $u(x,t) \in L^2((0,T)\times\mathbb{R}^d)$, 使得 u^ε 双尺度收敛于 $u(x,t)$. 此外, 估计

$$\|\nabla u^\varepsilon\|_{L^2((0,T)\times\mathbb{R}^d)} \leqslant C\varepsilon^{-1/2}$$

暗含了 $\varepsilon\nabla u^\varepsilon$ 在 $L^2((0,T)\times\mathbb{R}^d)$ 上弱收敛于 0.

接下来, 设 $\phi^\varepsilon(x,t) \in C_0^\infty(\mathbb{R}\times\mathbb{R}^d)$, (21.2.1) 的弱形式

$$\begin{aligned}&\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi^\varepsilon}{\partial t} - b^\varepsilon\cdot\nabla\phi^\varepsilon\right)u^\varepsilon\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}g(x)\phi^\varepsilon(x,0)\mathrm{d}x\\ &= \varepsilon D\int_{\mathbb{R}^+\times\mathbb{R}^d}\nabla u^\varepsilon\cdot\nabla\phi^\varepsilon\mathrm{d}x\mathrm{d}t.\end{aligned} \tag{21.3.2}$$

选择一个光滑的测试函数, 它独立于振荡 $\phi^\varepsilon = \phi(x,t)$. 在 (21.3.2) 中使用这个函数, 当 $\varepsilon \to 0$ 时通过极限. 使用 u^ε 的双尺度极限独立于 y 可得

$$\begin{aligned}&\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi^\varepsilon}{\partial t} - b^\varepsilon\cdot\nabla\phi^\varepsilon\right)u^\varepsilon\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}g(x)\phi^\varepsilon(x,0)\mathrm{d}x\\ &\to \int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi}{\partial t} - \bar{b}\cdot\nabla\phi\right)u\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}g(x)\phi(x,0)\mathrm{d}x,\end{aligned}$$

其中,

$$\bar{b} := \int_{\mathbb{T}^d}b(y)\mathrm{d}y.$$

此外, 由 $\varepsilon\nabla u^\varepsilon$ 在 $L^2((0,T)\times\mathbb{R}^d)$ 上弱收敛于 0 可得

$$\varepsilon D\int_{\mathbb{R}^+\times\mathbb{R}^d}\nabla u^\varepsilon\cdot\nabla\phi^\varepsilon\mathrm{d}x\mathrm{d}t \to 0.$$

结合前面的两个极限可得

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi}{\partial t}-\bar{b}\cdot\nabla\phi\right)u\mathrm{d}x\mathrm{d}t+\int_{\mathbb{R}^d}g(x)\phi(x,0)\mathrm{d}x=0,$$

这导致了 (21.2.2) 的弱形式 (见注 7.28). □

21.4 证明：$D=0$

定理 21.2 的证明 由定理 7.27 可知存在 (21.2.3) 的唯一弱解 u^ε 满足 $u^\varepsilon\in L^2((0,T);H^1(\mathbb{R}^d))$. 此外, 估计 (7.4.3) 给出

$$\|u^\varepsilon\|^2_{L^2(\mathbb{R}^+;L^2(\mathbb{R}^d))}\leqslant\|g\|^2_{L^2(\mathbb{R}^d)}\leqslant C. \tag{21.4.1}$$

由定理 2.38, 这个估计暗含了存在一个子序列, 仍用 u^ε 表示, 双尺度收敛于一个函数 $u_0\in L^2(\mathbb{R}^+\times\mathbb{R}^d;L^2_{\mathrm{per}}(\mathbb{T}^d))$.

由于 b 是无散度的, 对于每一个 $\phi^\varepsilon(x,t)\in C_0^\infty(\mathbb{R}\times\mathbb{R}^d)$, (21.2.3) 的弱形式, 方程 (7.4.9) 给出

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\left(\frac{\partial\phi^\varepsilon}{\partial t}-b^\varepsilon\cdot\nabla\phi^\varepsilon\right)u^\varepsilon\mathrm{d}x\mathrm{d}t+\int_{\mathbb{R}^d}g(x)\phi^\varepsilon(x,0)\mathrm{d}x=0, \tag{21.4.2}$$

选择如下形式的一个测试函数:

$$\phi^\varepsilon=\varepsilon\phi\left(x,\frac{x}{\varepsilon},t\right),\quad \phi\in C_0^\infty(\mathbb{R}\times\mathbb{R}^d;C^\infty_{\mathrm{per}}(\mathbb{T}^d)).$$

将 ϕ^ε 插入 (21.4.2), 运用链式法则 (ϕ^ε 是 x 和 x/ε 的函数), 当 ε 趋向于 0 时, 通过极限可得

$$\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\int_{\mathbb{T}^d}b(y)\cdot\nabla_y\phi(x,y,t)u_0(x,y,t)\mathrm{d}y\mathrm{d}x\mathrm{d}t=0.$$

这是方程 (21.2.4a). 选择一个测试函数

$$\phi^\varepsilon=\phi\left(x,\frac{x}{\varepsilon},t\right),\quad \phi\in C_0^\infty(\mathbb{R}\times\mathbb{R}^d;C^\infty_{\mathrm{per}}(\mathbb{T}^d)),$$

使得满足 (21.2.4b). 将这个函数插入 (21.4.2), 当 ε 趋向于 0 时, 通过极限可得

$$\begin{aligned}\int_{\mathbb{R}^+}\int_{\mathbb{R}^d}\int_{\mathbb{T}^d}&\left(\frac{\partial\phi(x,y,t)}{\partial t}-b(y)\cdot\nabla_x\phi(x,y,t)\right)u_0(x,y,t)\mathrm{d}y\mathrm{d}x\mathrm{d}t\\&+\int_{\mathbb{R}^d}g(x)\left(\int_Y\phi(x,y,0)\mathrm{d}y\right)\mathrm{d}x=0,\end{aligned}$$

确切地为 (21.2.4c). □

当然, 为了使方程的均匀化系统有意义, 需要证明它有唯一解. 这是如下定理的内容:

定理 21.4 方程组 (21.2.4) 存在唯一解 $u_0(t,x,y) \in L^2(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{T}^d)$.

证明 解的存在性遵从 u^ε 的双尺度极限的存在性. 下面证明其唯一性, 使用定理 7.27 中解的唯一性证明相同的参数. 设 $u_1(x,y,t)$, $u_2(x,x,t)$ 是具有相同初始条件的均匀化系统的两个解, 作差值

$$e(x,y,t) = u_1(x,y,t) - u_2(x,y,t).$$

由线性关系知这个函数满足与方程 (21.2.4) 一样的系统, 并且有零初始条件, 则有

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{T}^d} \left(\frac{\partial \phi(x,y,t)}{\partial t} - b(y) \cdot \nabla_x \phi(x,y,t) \right) e(x,y,t) \mathrm{d}y \mathrm{d}x \mathrm{d}t = 0,$$

其中 $T > 0$ 是任意却又固定的. 定理 7.27 的变式说明了差值 e 满足 $e \in L^2((0,T); H^1(\mathbb{R}^d); L^2(\mathbb{T}^d))$, 其中 $\dfrac{\partial e}{\partial t} \in L^2((0,T) \times \mathbb{R}^d \times \mathbb{T}^d)$. 因此, 可以将它作为先前方程的一个测试函数来推导,

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{T}^d} \left(\frac{\partial e(x,y,t)}{\partial t} - b(y) \cdot \nabla_x e(x,y,t) \right) e(x,y,t) \mathrm{d}y \mathrm{d}x \mathrm{d}t = 0.$$

暂时设 e 有紧支集, 对方程给出的第二项进行分部积分, 并利用 b 是无散度可得

$$\begin{aligned} &\int_{\mathbb{R}^d} \int_{\mathbb{T}^d} b(y) \cdot \nabla_x e(x,y,t) e(x,y,t) \mathrm{d}y \mathrm{d}x \\ =& -\int_{\mathbb{R}^d} \int_{\mathbb{T}^d} (b(y) e(x,y,t)) \cdot \nabla_x e(x,y,t) \mathrm{d}y \mathrm{d}x, \end{aligned}$$

从上式推导出这一项消失了. 一个近似变元暗含了相同的结论对 $e \in L^2((0,T); H^1(\mathbb{R}^d) \times L^2(\mathbb{T}^d))$ 是有效的, 从而对任意 $T > 0$ 可推导出

$$\frac{1}{2} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d} \int_{\mathbb{T}^d} |e(x,y,t)|^2 \mathrm{d}y \mathrm{d}x \mathrm{d}t = 0.$$

由于 e 初始为零, 这个方程遵循对于每一个 $t > 0$ 有

$$\int_{\mathbb{R}^d} \int_{\mathbb{T}^d} |e(x,y,t)|^2 \mathrm{d}y \mathrm{d}x = 0,$$

所以 $e(x,y,t) = 0$, 从而解 $u_0(x,y,t)$ 是唯一的. □

推论 21.3 的证明 $b(y)$ 是遍历性的, 无散度向量场情况下, 当且仅当 u_0 独立于 y 时满足方程 (21.2.4a). 这是因为分部积分说明这个方程是首阶偏微分方程 $\mathcal{L}_0 u_0 = 0$ 的弱形式, 其中 $\mathcal{L}_0 = -b(y) \cdot \nabla_y$ 在 $\mathbb{T}^d$ 上具有周期边界条件, 则通过选择一个独立于 y 的测度函数可以从 (21.2.4c) 中得到 (弱形式) 平均化方程 (21.2.5).□

21.5 讨论和参考

类似于定理 21.1 的结论可见文献 [5, 203, 253]; [203, 3.2 节] 是偏微分方程黏度解中基于扰动测试函数方法[96] 的一个证明. 当速度场 b 显式地依赖于 x 和 x/ε 时, 如下结论成立:

$$b^\varepsilon = b\left(x, \frac{x}{\varepsilon}\right),$$

其中 $b(x,y)$ 在 y 上是 1 周期的 (见练习第 5 题). 接着平均化速度场是 x 的一个函数, $\bar{b} = \bar{b}(x)$, 得到 (21.2.2) 的下一阶数的校正也是可能的[45,253]. 实际上, 函数 $\bar{u}^\varepsilon = u + \varepsilon\langle u_1\rangle$ 与 $O(\varepsilon^2)$ 同阶, 其中 $\langle\cdot\rangle$ 表示 $\mathbb{T}^d$ 上的平均化, 满足对流扩散方程

$$\frac{\partial \bar{u}^\varepsilon}{\partial t} + \bar{b}(x)\cdot\nabla\bar{u}^\varepsilon = \varepsilon\nabla\cdot(\mathcal{K}(x)\nabla\bar{u}^\varepsilon)$$

有效扩散率 $\mathcal{K}(x)$ 是通过一个适当单胞问题的解得到的. 注意到它是 x 的函数.

定理 21.2 的证明遵循文献 [80]. 在一般的二维情况下, 测试函数上的一个完全集 (在 L^2 上) 可以用来表征平均化 (有时是均匀化) 极限, 见文献 [147]. 这说明了此极限是线性双曲型方程的无限对称集. 特征法证明二维流的极限定理见文献 [312]. 对形式为

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\nabla\left(F(x) + \varepsilon V\left(\frac{x}{\varepsilon}\right)\right)$$

的常微分方程平均化问题 (以及相关的 Liouville 方程) 的考虑见文献 [226].

21.6 练　习

1. 考虑如下 Cauchy 问题:

$$\frac{\partial u^\varepsilon(x,t)}{\partial t} - b\left(x, \frac{x}{\varepsilon}\right)\cdot\nabla u^\varepsilon(x,t) = 0, \quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+,$$

$$u^\varepsilon(x,0) = g(0), \quad x\in\mathbb{R}^d,$$

其中 $b\in C_b^\infty(\mathbb{R}^d, C_{\mathrm{per}}^\infty(\mathbb{T}^d);\mathbb{R}^d)$ 且 $\nabla\cdot b(x, x/\varepsilon) = 0$. 运用双尺度收敛方法证明平均化定理.

2. 对受迫运输型偏微分方程

$$\frac{\partial u^\varepsilon(x,t)}{\partial t} - b(x)\cdot\nabla u^\varepsilon(x,t) = g\left(x, \frac{x}{\varepsilon}\right), \quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+,$$

$$u^\varepsilon(x,0) = g(x), \quad x\in\mathbb{R}^d$$

进行与第 1 题中相同的程序, 其中 $g(x,y)$ 在第二个变量中是光滑的、周期的, 并且 $g(x,x/\varepsilon)$ 在 $L^2(\mathbb{R}^d)$ 上有界.

3. 对具有振荡初始数据

$$u^{\varepsilon}(x,0)=f\left(x,\frac{x}{\varepsilon}\right)$$

的运输方程 (21.2.3a) 进行和第 2 题中相同的程序, 其中 $g(x,y)$ 在其二元中是光滑的、周期的, 并且 $g(x,x/\varepsilon)$ 在 $L^2(\mathbb{R}^d)$ 上有界.

4. 结合第 3 题证明运输偏微分方程

$$\frac{\partial u^{\varepsilon}(x,t)}{\partial t}-b\left(x,\frac{x}{\varepsilon}\right)\cdot\nabla u^{\varepsilon}(x,t)=g\left(x,\frac{x}{\varepsilon}\right),\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+,$$

$$u^{\varepsilon}(x,0)=f\left(x,\frac{x}{\varepsilon}\right),\quad x\in\mathbb{R}^d$$

的平均化定理.

5. 证明对流扩散方程

$$\frac{\partial u^{\varepsilon}(x,t)}{\partial t}-b\left(x,\frac{x}{\varepsilon}\right)\cdot\nabla u^{\varepsilon}(x,t)=\varepsilon D\Delta u^{\varepsilon}+g\left(x,\frac{x}{\varepsilon}\right),\quad (x,t)\in\mathbb{R}^d\times\mathbb{R}^+,$$

$$u^{\varepsilon}(x,0)=f\left(x,\frac{x}{\varepsilon}\right),\quad x\in\mathbb{R}^d$$

在第 4 题假设下的平均化定理.

参考文献

[1] A.Abdulle. Multiscale methods for advection-diffusion problems. *Discrete Contin. Dyn. Syst.*, suppl: 11-21, 2005.

[2] A.Abdulle and W.E.Finite difference heterogeneous multi-scale method for homogenization-n problems. *J.Comput. Phys.*, 191(1): 18-39, 2003.

[3] A.Abdulle and Ch.Schwab. Heterogeneous multiscale FEM for diffusion problems on rough surfaces. *Multiscale Model. Simul.*, 3(1): 195-220(electronic), 2004/05.

[4] R.A.Adams and J.J.F.Fournier. *Sobolev Spaces.* Pure and Applied Mathematics. Elsevier, Oxford, 2003.

[5] G.Allaire. Homogenization et convergence a deux echelles. application a un probleme de covection diffusion. *C.R.Acad. Sci. Paris*, 312: 581-6, 1991.

[6] G.Allaire. Homogenization and two-cale convergence. *SIAM J.Math. Anal.*, 23(6): 1482-1518, 1992.

[7] G.Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Math-ematical Sciences.* Springer-Verlag, New York, 2002.

[8] G.Allaire and M.Briane. Multiscale convergence and reiterated homogenization. Proc. Roy. Soc. Edinburgh sect. A, 126(2): 297-342, 1996.

[9] G.Allaire and R.Brizzi. A multiscale finite element method for numerical homogenization. *Multiscale Model. Simul.*, 4(3): 790-812(electronic), 2005.

[10] G.Allaire and A.Piatnitski. Homogenization of the Schrödinger equation and effective mass theorems. *Comm. Math. Phys.*, 258: 1-22, 2005.

[11] R.Aris On the dispersion of a solute in a fluid flowing through a tube. *Proc. R. Soc. Lond.* A, 235: 67-77, 1956.

[12] L. Arnold. *Stochastic differential equations*: theory and applications. Wiley-Intersc-ience [John Wiley & Sons], New York, 1974. Translated from the German.

[13] V.I.Arnol'd *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989.

[14] V.I.Arnol'd and A.Avez. *Erogodic problems of classical mechanics.* Translated from the French by A. Avez. W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[15] V.I.Arnol'd. *Ordinary differential equations.* Springer-Verlag, Berlin, 1992.

[16] G.Ben Arous and H.Owhadi. Multiscale homogenization with bounded ratios and anomalous slow diffusion. *Comm. Pure Appl. Math.*, 56(1): 80-113, 2003.

[17] Z.Artstein. On singularly perturbed ordinary differential equations with measure-valued limits. Math. Bohem., 127(2): 139-52, 2002.

[18] Z.Atstein and M.Slemrod. On singularly perturbed retarded functional-differential equations. J. *Differential Equations,* 171(1): 88-109, 2001.

[19] Z.Atrstein and A.Vigodner. Singularly perturbed ordinary differential equations with dynamic limits. *Proc. Roy. Soc. Edinburgh Sect.* A, 126(3): 541-69, 1996.

[20] M.Avellanada and A.J.Majda. Stieltjes integral representation and effective diffusivity bounds for turbulent transport. *Phys. Rev. Lett.*, 62: 753-5, 1989.

[21] M.Avellanada and A.J.Majda. An integral representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Comm. In Math. Phys.*, 1388: 339-91, 1991.

[22] M.Avellanada and M.Vergassola. Stieltjes integral representation of effective diffusivities in time-dependent flows. *Phys. Rev. E,* 52: 3245-51, 1995.

[23] I.Babuška. Homogenization and its application. Mathematical and computational problems. *In Numerical solution of partial differential equations,* III*(Proc.Third Sympos.(SYNSPAD-E), Univ. Maryland, College Park, Md.,* 1975*)*, Pages 89-116. Academic Press, New York, 1976.

[24] I.Babuška. Homogenization approach in engineering. In *Computing methods in applied sciences and engineering(Second Internat. Syspos., Versailles,* 1975*),* Part 1, pages 137-53. Lecture Notes in Econom. And Math. Systems, Vol. 134. Springer, Berlin, 1976.

[25] I.Babuška. Solution of interface problems by homogenization. I. *SIAM J. Matj. Anal.*, 7(5): 603-34, 1976.

[26] I.Babuška. Solution of interface problems by homogenization. I. *SIAM J. Matj. Anal.*, 7(5): 635-45, 1976.

[27] N.Bakhvalov and G.Panasenko. *Homogenization: averaging processes in periodic media.* Kluwer Academic Publishers Group, Dordrecht, 1989. Mathematical problems in the mechanics of composite materials. Translated from the Russian by D. Leites.

[28] R.Balescu. Statistical dynamics. *Matter out of equilibrium.* Imperial College Press, London. 1997.

[29] J.M.Ball. A version of the fundamental theorem for Young measures. In PDFs *and continuum models of phase transitions(Nice,* 1988*),* Volume 334 of Lecture Notes in Phys., pages 207-15. Springer, Berlin, 1989.

[30] D.Barkley, I.G.Kevrekidis, and A.M.Stuart. The moment map: nonlinear dynamics of density evolution via a few moments. SIADS, 5(3): 403-34, 2006.

[31] C.Beck. Brownian motion from deterministic dynamics. *Phys. A.* 169(2): 324-36, 1990.

[32] G.Benettin, L.Galgani and A.Giorgilli. Realization of homonymic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. I. *Comm. Math. Phys.*, 113(1)87-103, 1987.

[33] A.Bensoussan, J.-L.Lions and G.Papanicolaou. *Asymptotic analysis for periodic structures,* volume 5 of Studies in Mathematics and Its Applications. North-Holland

Publishing Co., Amsterdam, 1978.

[34] N.Berglund and B.Gentz. Noise-induced phenomena in slow-fast dynamical systems. Probability and Its Applications(New York). Springer-Verlag, London Ltd., London, 2006. A sample-paths approach.

[35] R.N.Bhattacharya, V.K.Gupata, and H.F.Walker. Asymptotic of solute dispersion in periodic porous media. SIAM J. APPL. MATH, 49(1): 86-98, 1989.

[36] P.Billingsley. Probability and measure. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[37] P.Billingsley. Convergence of probability measures. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1999.

[38] D.Blömker. Amplitude equations for locally cubic non-autonomous nonlinearities. SIAM Journal on Applied Dynamical Systems, 2(3): 464-86, 2003.

[39] D.Blömker. Approximation of the stochastic Rayleigh-Bénard problem near the onset of convection and related problems. Stoch. Dyn., 5(3): 441-74, 2005.

[40] D.Blömker. M.Hairer, and G.A.Pavliotis. Multiscale analysis for stochastic PDEs with quadratic nonlinearities. Nonlinearity, 20: 1721-44, 2007.

[41] D.Blömker. S.Maier-Paape, and G.Schneider. The stochastic Landau equation as an amplitude equation. Discrete and Continuous Dynamical Systems, Series B, 1(4): 527- 41, 2001.

[42] J.Bonn and R.M.Mclaughlin. Sensitive enhanced diffusivities for flows with fluctuating mean winds: a two- parameter study. J. Fluid Mech., 445: 345-75, 2001.

[43] F.Bornmann. Homogenization in time of singularly perturbed mechanical systems, volume 1687 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1998.

[44] F.A.Bornemann and Ch. Schüte. Homogenization of Hamiltonian systems with a strong constraining potential. Phys. D, 102(1-2).

[45] A.Bourgeat. M.Jurak and A.L.Piatnitski. Averaging a transport equation with small diffusion and oscillating velocity. *The Mathematics Preprint Server,* 2001.

[46] S.Brahim-Otsmane. G.A.Francfort and F.Murat. Homogenization in thermo elasticity. In *Random media and composites (Leesburg. VA. 1988),* pages 13-45. SIAM. Philadelphia, 1989.

[47] S.Brahim-Otsmane. G.A.Francfort and F.Murat. Correctors for the homogenization of the wave and heat equations. *J. Math. Pure Appl. (9),* 71(3): 197-231,1992.

[48] A.Braides. *Γ-convergence for beginners,* volume 22 of *Oxford Lecture Series in Mathematics and lts Application- ns.* Oxford University Press, Oxford, 2002.

[49] A.Braides and A.Defrancesehi. *Homogenization of multiple integrals,* volume 12 of *Oxford Lecture Series in Mathematics and Its Applications.* The Clarendon Press Oxford University Press, New York, 1998.

[50] L.Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM). Philadelphia, 1992. Corrected reprint of the 1968 original.

[51] Y.Brenier. Remarks on some linear hyperbolic equations with oscillatory coefficients. In *Third International Conference on Hyperbolic Problems, Vol.* I, I I *(Uppsala, 1990)*, pages 119-30. Studentlitteratur. Lund, 1991.

[52] M.E.Brewster and G.Beylkin. A multiresolution strategy for numerical homogenization. *Appl. Comput. Harmon. Anal.*, 2(4): 327-49, 1995.

[53] H.Brezis. *Analyse fonctionnelle.* Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983.

[54] F.Campillo and A.Piatnitski. Effective diffusion in vanishing viscosity. In *Nonlinear partial differential equations and their applications. Collège de France Seminar. Vol. XIV (Paris,*1997/1998), volume 31 of *Stud. Math. Appl.*, pages 133-45. North-Hol land, Amsterdam, 2002.

[55] R.A.Carmona and F.Cerou. Transport by incompressible random velocity fields: simulations and mathematical conjectures. In *Stochastic partial differential equations: six perspectives*, volume 64 of *Math. Surveys Monogr.*, pages 153-81. Amer. Math. Soc., Providence. RI. 1999.

[56] R.A.Carmona and L.*Xu*. Homogenization theory for time-dependent two-dimensional incompressible gaussian flows. *The Annals of Applied Probability*, 7(1): 265-79, 1997.

[57] J.Carr. *Applications of centre manifold theory*, volume 35 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1981.

[58] S.Ccrrai *Second order PDE's in finite and infinite dimension*, volume 1762 *of Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2001.

[59] X.Chao and A.J.Roberts. On the low-dimensional modeling of Stratonovich stochastic differential equations. *Physica A*, 225: 62-80, 1996. http://dx.doi.org/10.1016/0378-4371(95)00397-8.

[60] Z.Chen and T Y.Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Math. Comp.*, 72(242): 541-76 (electronic), 2003.

[61] A.Chertock and D.Levy. On wavelet-based numerical homogenization. *Multiscale Model. Simul.*, 3(1): 65-88 (electronic), 2004/05.

[62] S.Childress. Alpha effect in flux ropes and sheets. *Phys. Earth and Planet. Int.*, 20: 172-80. 1979.

[63] S.Childress and A.M.Soward. Scalar transport and alpha-effect for a family of cat's-eye flows. *J. Fluid Meck.*, 205: 99-133, 1989.

[64] A.J.Chorin and O.H.Hald. *Stochastic tools in mathematics and science*, volume 1 of *Surveys and Tutorials in the Applied Mathematical Sciences.* Springer, New York,

2006.

[65] A.J.Chorin. O.H.Hald and R.Kupferman. Optimal prediction with memory. *Phys. D,* 166(3-4): 239-57. 2002.

[66] D.Cioranescu and P.Donato. *An Introduction to Homogenization.* Oxford University Press, New York, 1999.

[67] D.Cioranescu and J.Saint Jean Pauline. *Homogenization of reticulated structures,* volume 136 of *Applied Mathematical Sciences.* Springer-Verlag, New York. 1999.

[68] E.A.Coddington and N.Levinson. *Theory of ordinary differential equations.* McGraw-Hill Book Company, Inc., New York. Toronto. London, 1955.

[69] P.Constantin, C.Foias, B.Nicolaenko and R.Temam. *Integral manifolds and inertia! manifolds for dissipative partial differential equations,* volume 70 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1989.

[70] I.P.Cornfeld, S.V.Fomin and Ya.G.Sinai. *Ergodic theory,* volume 245 of *Grunilehren der Mathemaiischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Vertag. New York. 1982. Translated from the Russian by A. B. Sosinskii.

[71] S.M.Cox and A.J.Roberts. Initial conditions for models of dynamical systems. *Physica* D, 85: 126-41, 1995.

[72] B.Dacorogna. *Direct methods in the calculus of variations.* Springer-Verlag, Berlin, Heidelberg. 1989.

[73] E.De Giorgi and S.Spagnolo. Sulla convergenza degli integral! dell'energia per operaiori cllittici del secondo ordine. *Boll. Un. Mat. Ital. (4),* 8: 391-411, 1973.

[74] A.De Masi, P. A. Ferrari, S. Goldstein, and W. D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Statist. Phys.,* 55(3-4):787-855, 1989.

[75] A.Defranceschi. An introduction to homogenization and G-convergence. Lecture notes. http://www.sissa. it/fa/publications/lecture_notes/, 1993.

[76] M.Dorobantu and B.Engquist. Wavelet-based numerical homogenization. *SIAM J. Numer. Anal.,* 35 (2): 540-59 (electronic). 1998.

[77] J.Duan, K.Lu and B.Schmalfuss. Invariant manifolds for stochastic partial differential equations. *Annals of Probability,* 31: 2109-35, 2006.

[78] J.Duan and B.Nadiga. Stochastic parameterization for large eddy simulation of geophysical flows. *Proc. Ameri. Math. Soc,.* 2006.

[79] E.B.Dynkin. *Markov processes. Vols.* I, II. volume 122 of *Translated with the authorization and assistance of the author by J. Fabius. V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wi ssenschaften. Bände* 121. Academic Press Inc., New York, 1965.

[80] W.E.Homogenization of linear and nonlinear transport equations. *Communications in Pure and Applied Mathematics.* 45: 301-26, 1992.

[81] W.E and B.Engquist. The heterogeneous muitiscale methods. *Commun. Math Sci.,* 1(1): 87-132, 2003.

[82] W.E and B.Engquisl. Mutiscale modeling and computation. *Notices Amer. Math. Soc.,* 50(9): 1062-70, 2003.

[83] W.E.B.Engquist. X.Li, W.Ren and E.Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Communications in Computational Physics,* 2(3): 367-450, 2007.

[84] W.E.D.Liu and E.Vunden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.,* 58(11): 1544-85, 2005.

[85] W.E.D.Liu and E.Vanden-Eijnden. Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J. Chem. Phys.,* 123: 194107, 2005.

[86] W.E.D.Liu and E.Vanden-Eijnden. Nested stochastic simulation algorithm for chemical kinetic systems with multiple time scales,. *J. Comp. Phys.,* 221(1): 158-180, 2007.

[87] R.E.Edwards. *Functional analysis.* Dover Publications Inc., New York. 1995.

[88] Y.Efendiev, T.Hou and V.Ginting. Multiscale finite element methods for nonlinear problems and their applications. *Commun. Math. Sci.,* 2(4): 553-89, 2004.

[89] Y.Efendiev and A.Pankov. Numerical homogenization of monotone elliptic operators. *Multiscale Model. Simul.,* 2 (1): 62-79 (electronic), 2003.

[90] Y.Efendiev and A.Pankov. Numerical homogenization and correctors for nonlinear elliptic equations. *SIAM J. Appl. Math.,.* 65(1): 43-68 (electronic). 2004.

[91] H.I.Ene and E.Sánchez-Palencia. Équations et phénomènes de surface pour I'écoulement dans un modèle de milieu poreux. *J. Mécanique.* 14: 73-108, 1975.

[92] B.Engquist. Wavelet based numerical homogenization. In *Proceedings of the International Congress of Mathematicians. Vol.* I I *(Berlin.* 1998*),* number Extra Vol. III, pages 503-512 (electronic). 1998.

[93] B.Engquist and Y.-H.Tsai. The multiscale heterogeneou methods for stiff ODEs. *Math. Comp.,* 74: 1707-42, 2005.

[94] S.N.Ethier and T.G.Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[95] L.C.Evans. *Weak convergence methods for nonlinear partial differential equations.* volume 74 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1990.

[96] L.C.Evans. The perturbed test function method for viscosity solutions of nonlinear pde. *Proc. Roy. Soc. Edinburgh,* 111A: 359-75, 1989.

[97] L.C.Evans. Periodic homogenization of certain fully nonlinear partial differential equations. *Proc, Roy. Soc. Edinburgh,* 120: 245-65, 1992.

[98] L.C.Evans. *Partial Differential Equations.* AMS, Providence. RI. 1998.

[99] G.Falkovich. K.Gawedzki and M.Vergassola. Particles and fields in fluid turbulence. *Rev. Modern Phys.,* 73(4): 913-75, 2001.

[100] A.Eannjiang and G.C.Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM J. APPL. MATH,* 54: 333-408, 1994.

[101] A.Eannjiang and G.C.Papanicolaou. Convection enhanced diffusion for random flows. *J. Stat. Phys.,* 88: 1033-76, 1997.

[102] W.Feller. *An introduction to probability theory and its applications. Vol.* I. Third edition. John Wiley & Sons Inc., New York, 1968.

[103] W.Feller. *An introduction to probability theory and its applications. Vol.* I I. Second edition. John Wiley & Sons Inc., New York, 1971.

[104] N.Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.,* 21: 193-226, 1971/1972.

[105] N.Fenichel. Geometric singular perturbation theory for ordinary differential equations. J. *Differential Equations,* 31(1): 53-98, 1979.

[106] J-P.Fouque, G.C.Papanicolaou and R.K.Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge, 2000.

[107] G.A.Francfort and F.Murat. Homogenization and optimal bounds in linear elasticity. *Arch. Rational Mech. Anal.,* 94(4): 307-34, 1986.

[108] G.A.Francfort and P-M.Suquet. Homogenization and mechanical dissipation in thermoviscoelasticity. *Arch. Rational Mech. Anal.,* 96(3): 265-93, 1986.

[109] M.Freidlin. Some remarks on the Smoluchowski-Kramers approximation. J. *Statist. Phys.,* 117(3-4): 617-34, 2004.

[110] M.I.Freidin. The Dirichletl prohlem for an equation with periodic coefficients depending on a small parameter. *Teor. Verajatnost. i Primenen.,* 9: 133-9, 1964.

[111] M.I.Freidlin and A.D.Wemzell. *Random perturbations of dynamical systems.* Springer-Verlag, New York, 1984.

[112] A.Friedman. *Partial differential equations of parabolic type.* Prentice-Hall Inc., Engle-wood Cliffs, NJ, 1964.

[113] A.Friedman. *Partial differential equations.* Holt, Rinehart and Winston, New York, 1969.

[114] A.Friedman. *Stochastic differential equations and applications. Vol.* 1. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. Probability and Mathematical Statistics, Vol. 28.

[115] F.R.Gantmacher. *The theory of matrices. Vol.* 1. AMS Chelsea Publishing, Providence. RI, 1998. Translated from the Russian by K. A. Hirsch, Reprint of the 1959

translation.

[116] B.García-Archilla, J.M.Sanz-Serna and R.D.Skeel. Long-time-step methods for oscillatory differential equations. *SIAM J. Sci. Comput.*, 20(3): 930-63 (electronic), 1999.

[117] C.W.Gardiner. *Handbook of stochastic methods.* Springer-Verlag, Berlin, second edition. 1985.

[118] J.Gamier. Homogenization in a periodic and lime-dependent potential. *SIAM J. APPL MATH.* 57(l): 95-lll, 1997.

[119] C.W.Gear and I.G.Kevrekidis. Projective methods tor stiff differential equations: problems with gaps in their eigenvalue spectrum. *SIAM J. Set. Compute* 24(4): 109l-1106 (electronic). 2003.

[120] I.I.Gikhman and A.V.Skorokhod. *Introduction to the theory of random processes.* Dover Publications Inc., Mineola, NY, 1996.

[121] D.Gilbarg and N.S.Trudinger. *Elliptic partial differential equation of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001.

[122] D.T.Gillespie. A general method for numerically simulating the stochastic lime evolution of coupled chemical reactions.*J. Computational Phys.*, 22(4): 403-34, 1976.

[123] D.Givon, I.G.Kevrekidis and R.Kupferman. Strong convergence schemes of projective integration schemes for singularly perturbed stochastic differential equations. *Comm. Math. Sci.*, 4(4): 707-29, 2006.

[124] D.Givon and R.Kupferman. White noise limits for discrete dynamical systems driven by fast deterministic dynamics. *Phys. A*, 335(3-4): 385-412, 2004.

[125] D.Givon, R.Kupfennan and A.M.Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity,* 17(6): R55-KI27, 2004.

[126] K.Golden. S.Goldstein and J.L.Lebowitz. Classical transport in modulated structures. *Phys. Rev. Lett.*, 55(24): 2629-32, 1985.

[127] S.Goldstein. Antisymmetric functional of reversible Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(1): 177-90, 1995.

[128] O.Gonzalez and A.M.Stuart. *Continuum Mechanics.* Texts in applied mathematics. Cambridge University Press. 2008.

[129] R.Graham and A.Schcnzle. Stabilization by multiplicative noise. PAvv. *Rev. A.* 26(3): 1676-85, 1982.

[130] G.R.Grimmett and D.R.Stirzaker. Probability and random processes. Oxford University Press, New York, 2001.

[131] J.Guckenheimer and P.Holmes. Nonlinear oscillations, dynamical systems, and bifurcations of vector fields, volume 42 of Applied Mathematical Sciences. Springer-Verlag. New York, 1983.

[132] R.B.Guenther and J.W.Lee. *Partial differential equations of mathematical physics and integral equations.* Dover Publications Inc., Mineola, NY, 1996.

[133] M.Briane. G.W.Milton and J.R.Willis. On cloaking for elasticity and physical equations with a trans-formation invariant form. *New J. Phys.,* 8: 248-67, 2006.

[134] M.Hairer and G.A.Pavliotis. Periodic homogenization for hypo elliptic diffusions. *J. Statist. Phys.,* 117(l-2): 261-79, 2004.

[135] J.K.Hale. Averaging methods for differential equations with retarded arguments and a small parameter. *J. Differential Equations,* 2: 57-73, 1966.

[136] J.K Hale. *Asymptotic behavior of dissipative systems,* volume 25 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI. 1988.

[137] P.Hall and C.C.Heyde. *Martingale limit theory and its application.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.

[138] P.Hanggi. P.Talkner and M.Borkovec. Reaction-rate theory: fifty years after Kramers. *Rev. Modern Phys.,* 62(2): 251-341. 1990.

[139] R.Z.Has'minskfi. *Stochastic stability of differential equations,* volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis.* Sijlhoff & Noordhoff. Alphcn aan den Rijn. 1980.

[140] S.Heinze. Diffusion-advection in cellular flows with large Peclet numbers. *Arch. Ration. Mech. Anal.,* 168(4): 329-42, 2003.

[141] V.H.Hoang and Ch.Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Model. Simul.,* 3(1): 168-94 (electronic). 2004/05.

[142] A.Holmbom. Homogenization of parabolic equations: An alternative approach and some corrector-type results. *Applications of Mathematics.* 42(5): 321-43, 1995.

[143] M.H.Holmes. *Introduction to perturbation methods.* Springer-Verlag, New York. 1995.

[144] W.Horsthemke and R.Lefever. *Noise-induced transitions,* volume 15 of *Springer Series in Synergetics.* Springer-Verlag, Berlin, 1984. Theory and applications in physics, chemistry, and biology.

[145] T.Y.Hou and X-H.Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.,* 134(1): 169-89, 1997.

[146] T.Y.Hou, X-H.Wu and Z.Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.,* 68(227): 913-43, 1999.

[147] T.Y.Hou and X.Xin. Homogenization of linear transport equations with oscillatory vector fields. *SIAM J. APPL. MATH,* 52(0): 34-45, 1992.

[148] W.Huisinga, Ch.Sehütte and A.M.Stuart. Extracting macroscopic stochastic dynamics: model problems. *Comm. Pure Appl. Math.,* 56(2): 234-69, 2003.

[149] G.Humer and I.G.Kevrekidis. Coarse molecular dynamics of *a* peptide fragment: free energy, kinetics and long lime dynamics computations. *J. Chem. Phys.*, 118(23): 1062-73, 2003.

[150] A.R.Humphries and A.M.Stuart. Deterministic and random dynamical systems: theory and numerics. In *Modern methods in scientific computing and applications (Montréal, QC,* 2001). volume 75 of *NATO Sci. Ser.* II *Math. Phys. Chan.*, pages 211-54. Kluwer Acad. Publ., Dordrecht. 2002.

[151] N.Ikeda and S.Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland Publishing Co., Amsterdam, second edition. 1989.

[152] J.Jacod and A.N.Shiryaev. *Limit theorems for stochastic processes,* volume 288 of *Crundlehren der Mathematischen Wissenschaflen* [*Fundamental Principles of Mathematical Sciences*]. Springer-Verlag, Berlin, 2003.

[153] V.V.Jikov, S.M.Kozlov and O.A.Oleinik. *Homogenization of differential operators and integral functionals.* Springer, Berlin, 1994.

[154] W.Just, K.Gelfert, N.Baba, A.Riegert and H.Kantz. Elimination of fast chaotic degrees of freedom: on the accuracy of the Born approximation. *J. Statist. Phys.*, 112(1-2): 277-92, 2003.

[155] W.Just, H.Kantz, *Ch.*Rödenbeck, and M. Helm. Stochastic modelling: replacing fast degrees of freedom by noise. *J. Phys. A,* 34(15): 3199-3213, 2001.

[156] I.Karatzas and S.E.Shreve. *Brownian motion and stochastic calculus,* volume 113 of *Graduate Tens in Mathematics.* Springer-Verlag. New York, second edition. 1991.

[157] T.Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag. Berlin, 1995. Reprint of the 1980 edition.

[158] J.B.Keller. Daley's law for flow in porous media and the two-space method. In *Nonlinear partial differential equations in engineering and applied science* (*Proc. Conf. Univ. Rhode Island. Kingston. R.I.,* 1979), volume 54 of *Lecture Notes in Pure and Appl. Math.,* pages 429-43. Dekker. New York, 1980.

[159] J.B.Keller. Effective behavior of heterogeneous media. In *Statistical mechanics and statistical methods in theory and application* (*Proc. Sympos., Univ. Rochester, Rochester, NY., 1976*), pages 631-14. Plenum, New York, 1977.

[160] J.B.Keller and J.Rubinstein. Nonlinear wave motion in a strong potential. *Wave Motion,* 13(3): 291-302, 1991.

[161] J.Kevorkian and J.D.Cole. *Multiple scale and singular perturbation methods,* volume 114 of *Applied Mathematical Sciences.* Springer-Verlag. New York. 1996.

[162] I.G.Kevrekidis, C.W.Gear, J.M.Hyman, R.G.Kevrekidis, O.Runborg and C.Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.,* 1(4): 715-62, 2003.

[163] I.G.Kevrekidis. A.Makeev and D.Maroudas. "coarse" stability and bifurcation analysis using timesteppers. *J. Chem. Phys.*, 116: 10083-91, 2002.

[164] I.G.Kevrekidis. Y.H.Qian and K.Theodoropoulos. "Coarse" stability and bifurcation analysis using stochastic simulators: Kinetic Monte Carlo examples. *Proc. Natl. Acad. Sci.*, 97(18): 9840-3, 2000.

[165] R.Z.Khasminskii. The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion. *Teor. Verojatnost. i Primenen.*, 8: 3-25-1963.

[166] R.Z.Khasminskii. A limit theorem for solutions of differential equations with a random right hand pan. *Teor. Verojatnost. i Primenen*, 11: 444-62, 1966.

[167] R.Z.Khasminskii and G.Yin. Asymptotic series for singularly perturbed Kolmogorov-Fokker-Planck equations. *SIAM J. Appl. Math.* 56(6): 1766-93, 1996.

[168] Y.Kifer. *Random perturbations of dynamical systems*, volume 16 of *Progress in Probability and Statistics* Birkhäuser Boston Inc., Boston. 1988.

[169] Y.Kifer. Averaging in dynamical systems and large deviations. *Invent. Math.*, 110(2): 337-70, 1992.

[170] Y.Kifer. Limit theorems in averaging for dynamical systems. *Ergodic Theory Dynam. Systems*, 15(6): 1143-72, 1995.

[171] Y.Kifer. Averaging and climate models. In *Stochastic climate models (Chorin.* 1999*)*. Volume 49 of *Progr. Probab.*, pages 171-88. Birkhäuser, Basel, 2001.

[172] Y.Kifer. Stochastic versions of Anosov's and Neistadt's theorems on averaging. *Stoch. Dyn.*, 1(1): 1-21, 2001.

[173] C.Kipnis and S.R.S.Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(l): l-19, 1986.

[174] D.L.Koch and J.F.Brady. The symmetry properties of the effective diffusivity tensor in anisotropic porous media. *Phys. Fluids*, 31(5): 965-73, 1988.

[175] D.L.Koch. R.G.Cox, H.Brenner and J.F.Brady. The effect of order on dispersion in porous media. *J. Fluid Mech.*, 200: 173-88, 1989.

[176] B.O.Koopman. Hamiltonian systems of transformations in Hilbert spaces. *Proc. Nat. Acad. Sri.* (*U.S.A*), 17: 315-18, 1931.

[177] S.M.Kozlov. The averaging of random operators. *Mat. Sb.* (*N.S.*), 109(151)(2): 188-202, 327, 1979.

[178] H-OKreiss. Problems with different time scales. In *Acta numerica, 1992.* Acta Numer., pages 101-39. Cambridge Univ. Press. Cambridge, 1992.

[179] E.Kreyszig. *Introductory functional analysis with applications.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1989.

[180] R.Kupferman. G.A.Pavliotis and A.M.Stuart. Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Phys. Rev. E*(3), 70(3): 036120, 2004.

[181] T.G.Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Functional Analysis*, 12: 55-67, 1973.

[182] T.G.Kurtz. Limit theorems and diffusion approximations for density dependent Markov chains. *Math. Programming Stud.*, 5: 67-78, 1976.

[183] T.G.Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3): 223-40, 1977/78.

[184] T.G.Kurtz. The central limit theorems for Markov chains. *The Annals of Probability*, 9: 557-60, 1981.

[185] C Landim. Central limit theorem for Markov processes. In *From classical to modern probability*, volume 54 *of Progr. Probab.*, pages 145-205. Birkhäuser, Basel, 2003.

[186] C.Landim and S.Olla. *Central limit theorems for Markov processes.* Lecture Notes, 2005.

[187] B.Lapeyre. E.Pardoux and R.Sentis. *Introduction to Monte-Carlo methods for transport and diffusion equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics.* Oxford University Press. Oxford, 2003. Translated from the 1998 French original by Alan Craig and Fionn Craig.

[188] A.Lasota and M.C.Mackey. *Chaos, fractals, noise.* Springer-Verlag, New York, 1994.

[189] P.D.Lax, *Functional analysis.* Wiley, NY., 2002.

[190] J.Li. P.G.Kevrekidis. C.W.Gear and I.G.Kevrekidis. Deciding the nature of the coarse equation through microscopic simulations: the baby-bathwater scheme. *Multiscale Model. Simul.*, 1(3): 391-407 (electronic), 2003.

[191] S.Lifson and J.L.Jackson. On the self-diffusion of ions in polyelectrolytic solution. *J. Chem. Phys.* 36: 2410, 1962.

[192] T.Lindvall. *Lectures on the coupling method.* Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.

[193] P.Lochak and C.Meunier. *Multiphase averaging for classical systems*, volume 72 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1988.

[194] M.Loeve. *Probability theory.* I. Springer-Verlag, New York, fourth edition, 1977. Graduate Texts in Mathematics, Vol. 45.

[195] M.Loeve. *Probability theory.* II. Springer-Verlag. New York, fourth edition, 1978. Graduate Texts in Mathematics, Vol. 46.

[196] L.Lorenzi and M.Bertoldi. *Analytical methods for Markov semigroups.* CRC Press, New York, 2006.

[197] D.Lukkassen, G.Nguetseng and P.Wall. Two-scale convergence. *Int. J. Pure Appl. Math.*. 2(1): 35-86-2002.

[198] T.J.Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana.* 14(2): 2l5-310, 1998.

[199] R.S.MacKay. Slow manifolds. In *Energy Localisation and Transfer,* pages 149-92. World Scientific, 2004.

[200] A.J.Majda and P.R.Kramer. Simplified models for turbulent diffusion: Theory, numerical modelling and physical phenomena. *Physics Reports.* 314: 237-574, 1999.

[201] A.J.Majda and R.M.McLaughlin. The effect of mean flows on enhanced diffusivity in transport by incompressible periodic velocity fields. *Stud. Appl. Math.,* 89431: 245-79, 1993.

[202] A.J.Majda. R.V.Abramov and M.J.Grote. *Information theory and stochastics for multiscale nonlinear systems,* volume 25 of *CRM Mongraph Series.* AMS, Providence, RI. 2005.

[203] A.J.Majda and P.Souganidis. The effect of turbulence on mixing in prototype reaction-diffusion systems. *Comm. Pure and Appl, Math.,* 53: 1284-1304, 2000.

[204] A.J.Majda and I.Timofeyev. Remarkable statistical behavior for truncated Burgers-Hopf dynamics. *Proc. Nati Acad. Sci. USA,* 97(23): 12413-17 (electronic), 2000.

[205] A.J.Majda, L Timofeyev and B.Vandcn Eijudcn. A mathematical framework lor stochastic climate models. *Comm. Purr. Appl Math.,* 54(8): 89l-974, 2001.

[206] A.J.Majda, I.Timofeyev and E.Vanden Eijnden. Models for stochastic climate prediction. *Proc. Natl. Acad. Sci. USA,* 96(26): 14687-91 (electronic), 1999.

[207] A.J.Majda. 1.Timofeyev and E.Vandcn-Eijndcn. Systematic strategies for stochastic mode reduction in climate. *J. Atmospheric Sci.,* 60(l4): 1705-22, 2003.

[208] A.M.Majda. I.Timofeyev and E.Vanden-Eijnden. A priori tests of a stochastic mode reduction strategy. *Phys. D,* 170(3-4): 206-52, 2002.

[209] A.M.Majda, I.Timofeyev and E.Vanden-Eijnden. Stochastic models for selected slow variables in large deterministic systems. *Nonlinearity,* 19(4): 769-94, 2006.

[210] X.Mao. *Stochastic differential equations and their applications.* Horwood Publishing Series in Mathematics & Applications. Horwood Publishing Limited, Chichester, 1997.

[211] X.Mao and C.Yuan. *Stochastic differential equations with Markovian switching.* Imperial College Press. London. 2006.

[212] J.E.Marsden and M.McCracken. *The Hopf bifurcation and its applications.* Springer-Verlag. New York, 1976. With contributions by P.Chernoff. G.Childs, S.Chow, J.R. Dorroh, J.Guckenheimer, L.Howard, N.Kopell, O.Lanford, J.Mallet-Paret, G.Oster, O.Ruiz, S.Sehecler, D.Schmidt and S.Smale, Applied Mathematical Sciences, Vol. 19.

[213] M.-L.Mascarenhas and A.-M.Toader. Scale convergence in homogenization-*Numer. Func.Anal. Optim,.* 22: 127-58, 2001.

[214] G.Dal Maso. *An Introduction to* $\Gamma-$*convergence.* Birkhauser. Boston. 1993.

[215] A.-M.Matache and Ch.Schwab. Finite dimensional approximations for elliptic problems with rapidly oscillating coefficients. In *Multiscale problems in science and technology (Dubrovnik,* 2000), pages 203-42. Springer, Berlin, 2002.

[216] A.-M.Matache and Ch.Schwab. Two-scale FF.M for homogenization problems. *M2-AN Math. Model. Numer. Anal,* 36(4): 537-72, 2002.

[217] M.R.Maxey and J.J.Riley. Equation of motion for a small rigid sphere in a nonuniform flow. *Phys. Fluids,* 26: 883-9, 1983.

[218] R.M.Mazo. *Brownian motion,* volume 112 *of International Series of Monographs on Physics.*Oxford University Press, New York, 2002.

[219] A.Mazzino. Effective correlation times in turbulent scalar transport. *Phys. Rev. E,* 56: 5500-10, 1997.

[220] A.Mazzino, S.Musacchio and A.Vulpiani. Multiple-scale analysis and renormalization for preasymptotic scalar transport. *Phys. Rev. E*(3), 71(1): 011113, 2005.

[221] D.W.McLaughlin. G.C.Papanicolaou and O.R.Pironneau. Convection of mitrostructure and related problems. *SIAMJ. APPL MATH,* 45: 780-97, 1985.

[222] R.M.McLaughlin. Numerical averaging and fast homognization. *J, Statist. Phys.,* 90(3-4): 597-626, 1998.

[223] R.M.McLaughlin and R.M.Forest. An anelastic. scale-separated model for mixing, with application to atmospheric transport phenomena. *Phys. Fluids,* 11(4): S80-92, 1999.

[224] I.Melbourne and M.Nicol. Large deviations for nonuniformly hyperbolic systems. Preprint, 2006.

[225] I.Melbourne and A.M.Stuart. SDE limits for scale-separated mixing ODEs. Preprint, 2007.

[226] G.Menon. Gradient systems with wiggly energies and related averaging problems. *Arch. Ration. Mech. Anal.,* 162(3): 193-246, 2002.

[227] S.P.Meyn and R.L.Tweedie. *Markov chains and stochastic stability.* Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[228] I.Mezic. J.F.Brady and S.Wiggins. Maximal effective diffusivity for time-periodic incompressible fluid flows. *SIAM J. APPL. MATH,* 56(I): 40-5fi, 1996.

[229] G.W.Milton. *The theory of composites,* volume 6 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press. Cambridge, 2002.

[230] G.W.Milton and J.R.Willis. On modifications of newton's second law kind linear continuum elastodynamics. *Proc. R Soc. A,* 463: 855-80, 2007.

[231] P.Ming and X.Yue. Numerical methods for multiscale elliptic problems. *J. Comput. Phys.,* 214(1): 421-15, 2006.

[232] F.Murat and L.Tartar. H-convergence. In *Topics in the mathematical modelling of composite materials,* volume 31 of *Progr. Nonlinear Differential Equations Appl.*. pages 21-43. Birkhäuser Boston, Boston, 1997.

[233] G.Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAMJ. Math. Anal.*, 20(3): 608-23, 1989.

[234] G.Nguetseng. Asymptotic analysis for a stiff variational problem arising in mechanics. *SIAMJ. Math. Anal.*, 21(6): 1394-1414, 1990.

[235] J.R.Norris, *Markov chains.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1998.

[236] B.ϕksendal. *Stochastic differential equations.* Universitext. Springer-Verlag, Berlin, 2003.

[237] O.A.Oleĭnik A.S.Shamacv and G.A.Yosifian. *Mathematical problems in elasticity and homogenization,* volume 26 of *Studies in Mathematics and* to *Applications.* North-Holland Publishing Co., Amsterdam, 1992.

[238] S.Olla *Homogenization of diffusion processes in random fields,* lecture Notes, 1994.

[239] R.E.O'Malley. Jr. *Singular perturbation methods for ordinary differential equations.* volume 89 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1991.

[240] G.C.Papanicolaou. Some probabilistic problems and methods in singular perturbations. *Rocky Mountain J. Math.*, 6(4): 653-74, 1976.

[241] G.C.Papanicolaou. Introduction to the asymptotic analysis of stochastic equations- In *Modern modeling of continuum phenomena* (*Ninth Summer Sem. Appl. Math., Rensselaer Polytech. Inst., Troy. NX.* 7975). pages 109-47. Lectures in Appl. Math., Vol. 16. Amer. Math. Soc., Providence, RI, 1977.

[242] G.C.Papanicolaou and W.Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Comm. Pure Appl. Math.*, 27: 641-68, 1974.

[243] G.C.Papanicolaou, D.W.Stroock and S.R.S.Varadhan. Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference* (*Duke Univ., Durham, N.C.,* 1976), *Paper No.* 6, pages ii+ 120 pp. Duke Univ. Math. Ser., Vol. 111. Duke Univ., Durham, NC, 1977.

[244] G.C.Papanicolaou and S.R.S.Varadhan. A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations. *Comm. Pure Appl. Math.*, 26: 497-524, 1973.

[245] G.C.Papanicolaou. Asymptotic analysis of transport processes. *Bull. Amer. Math. Soc.*, 81: 330-92, 1975.

[246] G.C.Papanicolaou. Asymptotic analysis of stochastic equations. In *Studies in probability theory,* volume 18 of MAA *Stud. Math.*, pages 111-79. Math. Assoc. America. Washington, DC, 1978.

[247] G.C.Papanicolaou and S.R.S.Varadhan. Ornstein-Uhlenbeck process in a random potential; *Comm. Pure Appl. Math.*, 38(6): 819-34, 1985.

[248] E.Pardoux. Homogenization of linear and semilinear second order parabolic pdes with periodic coefficients: A probabilistic approach. *Journal of Functional Analysis,* 167: 498-520, 1999.

[249] E.Pardoux and A.Yu. Veretennikov. On the Poisson equation and diffusion approximation. *I. Ann. Probab.*, 29(3): 1061-85, 2001.

[250] E.Pardoux and A.Yu. Veretennikov. On Poisson equation and diffusion approximation. II. *Ann. Probab.*, 31(3): 1166-92, 2003.

[251] E.Pardoux and A.Yu.Veretennikov. On the Poisson equation and diffusion approximation. III. *Ann. Probab.*, 33(3): 1111-33, 2005.

[252] K.R.Parthasarathy. *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence. RI. 2005. Reprint of the 1967 original.

[253] G.A.Pavliotis. *Homogenization theory for advection-diffusion equations with mean flow,* PhD Thesis. Rensselaer Polytechnic Institute, Troy, NY, 2002.

[254] G.A.Pavliotis and A.M.Stuart. White noise limits for inertial particles in a random field. *Multiscale Model. Simul.*, 1(4): 527-33 (electronic), 2003.

[255] G.A.Pavliotis and A.M.Stuart. Analysis of white noise limits for stochastic systems with two fast relaxation times. *Multiscale Model. Simul.*, 4(1): 1-35 (electronic), 2005.

[256] G.A.Pavliotis and A.M.Stuart. Periodic homogenization Tor inertial panicles. *Phys.*D, 204(3-4): 161-87, 2005.

[257] G.A.Pavliotis. A multiscale approach to Brownian motors. *Phys. Lett. A,* 344: 331-45, 2005.

[258] G.A.Pavliotis and A.M Stuart. Parameter estimation for multiscale diffusions. *J. Stat. Phys.*, 127(4): 741-81, 2007.

[259] A.Pazy. *Semigroups of linear operators and applications to partial differential equations.* Springer-Verlag. New York, 1983.

[260] K.Petersen. *Ergodic theory,* volume 2 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1983.

[261] M.A.Pinsky. *Lectures on random evolution.* World Scientific Publishing Co. Inc., River Edge, NJ, 1991.

[262] P.K.Pollen and A.J.Roberts. A description of the long-term behavior of absorbing continuous time Markov chains using a centre manifold. *Advances Applied Probability,* 22: 111-28, 1990.

[263] G.Da Prato and I.Zabczyk. *Stochastic equations in infinite dimensions,* volume 44 of *Encyclopedia of Mathematics and Its Applications.* Cambridge University Press, 1992.

[264] M.H.Protter and H.F.Weinberger. *Maximum principles in differential equations.* Springer-Verlag, New York, 1984.

[265] S.Raimondeau and D.G.Vlachos. Recent developments on multiscale, hierarchical modeling of chemical reactors. *Chem. Eng. J.*, 90: 3-23, 2002.

[266] M.Reed and B.Simon. *Methods of Modern Mathematical Physics I: Functional Analysis.* Academic Press, New York. 1980.

[267] P.Reimann. Brownian motors: noisy transport far from equilibrium. *Phys. Rep.,* 361(2-4): 57-265, 2002-

[268] P.Reimann. C.Van den Broeck, H.Linke, J.M.Rubi and A.Perez-Madrid. Giant acceleration of free diffusion by use of lilted periodic potentials. *Phys, Rev. Lett.,* 87(1): 010602, 2001.

[269] P.Resibois and M.De Leener. *Classical kinetic theory of fluids.* Wiley, New York, 1977.

[270] D.Revuz and M.Yor. *Continuous martingales and Brownian motion,* volume 293 of *Grundlehren der Mathematischen Wissenschuaften* [*Fundamental Principles of Mathematical Sciences*]. Springer-Verlag, Berlin, third edition, 1999.

[271] H.Risken. *The Fokker-Planck equation,* volume 18 of *Springer Series in Synergetics.* Springer-Verlag, Berlin, 1989.

[272] A.J.Roberts. Appropriate initial conditions for asymptotic descriptions of the long term evolution of dynamical systems.*J . Austrat. Math. Soc. B,* 31: 48-75, 1989.

[273] A.J.Roberts. Resolving the multitude of microscale interactions accurately models stochastic partial differential equations. *LMS J. Comput. Math.,* 9: 193-221, 2006.

[274] J.C.Robinson. *Infinite-dimensional dynamical systems.* Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2001.

[275] L.C.G.Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol.* I. Cambridge Mathematical Library. Cambridge University Press. Cambridge, 2000.

[276] L.C.G.Rogers and D.Williams. *Diffusions. Markov processes, anil martingales. Vol.* 2. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.

[277] H.Rubin and P.Ungar. Motion under a strong constraining force. *Comm. Pure Appl. Math.,* 10: 65-87, 1957.

[278] G.Samaey, D.Roose, and I.G.Kevrekidis. The gap-tooth scheme for homogenization problems. *Multiscale Model. Simul.,* 4(1): 278-306 (electronic), 2005.

[279] E.Sanchez-Palencia. *Nonhomogeneous media and vibration theory,* volume 127 of *lecture Notes in Physics.* Springer-Verlag, Berlin, 1980.

[280] J.M.Sancho, M.San Miguel and D.Dürr. Adiabatic elimination for systems of Brownian particles with nonconstant damping coefficients. *J. Statist. Phys.,* 28(2): 291-305, 1982.

[281] J.A.Sanders and F.Verhulst. *Averaging methods in nonlinear dynamical systems, volume* 59 *of Applied Mathematical Sciences.* Springer-Verlag, New York. 1985.

[282] M Schreier, P.Reimann. P.Hänggi and E.Pollak. Giant enhancement of diffusion and panicle selection in rocked periodic potentials. *Europhys. let.,* 44(4): 416-22, 1998.

[283] Z.Schuss. *Theory and applications of stochastic differential equations.* John Wiley & Sons Inc., New York. 1980, Wiley Series in Probability and Statistics.

[284] Ch.Schütte and W.Huisinga. Biomolecular conformations can be identified as metastable sets of molecular dynamics. In *Handbook of Numerical Analysis (Computational Chemistry), Vol. X,* 2003.

[285] Ch.Schütte. J.Walter, C.Hartmann and W.Huisinga. An averaging principle for fast degrees of freedom exhibiting long-term correlations. *Multiscale Model. S*imur., 2(3); 501-26 (electronic),2004.

[286] Ch.Schwab. Two scale FEM for homogenization problems. In *Mathematical modeling and numerical simulation in continuum mechanics (Yamaguchi.* 2000), volume 19 of *Lect. Notes Comput. Sci. Eng.,* pages 91 -107. Springer, Berlin. 2002.

[287] Ch.Schwab and A.-M.Matache. Generalized FEM for homogenization problems, in *Multiscale and multiresolution methods,* volume 20 of *Lect. Notes Comput. Sci. Eng.,* pages 197-237. Springer, Berlin, 2002.

[288] H.Sigurgeirsson and A.M.Stuart. Inertial particles in a random field. *Stoch. Dyn.,* 2(2): 295-310, 2002.

[289] H.Sigurgeirsson and A.M.Stuart. A model for preferential concentration. *Phys. Fluids,* 14(12): 4352-61, 2002.

[290] Ya.G.Sinai. *Introduction to ergodic theory.* Princeton University Press. Princeton. *NJ,* 1976. Translated by V. Scheffer, Mathematical Notes, 18.

[291] A.V.Skorokhod. *Asymptotic methods in the theory of stochastic differential equations.* volume 78 of *Translations of Mathematical Monographs.* American Mathematical Society. Providence, RI, 1989.

[292] T.H.Solomon and J.P.Gollub. Chaotic particle transport in time-dependent Rayleigh-Benard convection. *Phys. Fluids,* 38(12): 6280-86, 1988.

[293] T.H.Solomon and J.P.Gollub. Passive transport in steady Rayleigh-Benard convection. *Phys. Fluids*, 31(6): 1372-8, 1988.

[294] A.M.Soward. Fast dynamo action in a steady flow. *J. Fluid Meck.,* 180: 267-95, 1987.

[295] A.M.Soward and S.Childress-Large magnetic Reynolds number dynamo action in spatially periodic flow with mean motion. *Proc. Roy. Soc. Land, A,* 331: 649-733, 1990.

[296] S.Spagnolo. Sul limite delle soluzioni di problemi di Cauchy relativi all'equazione del calore. *Ann. Scuola Norm. Sup. Pisa* (3), 21: 657-99, 1967.

[297] S.Spagnolo. Sulla convergent di soluzioni di equazioni paraboliche ed elliltiche. *Ann. Scuola Norm. Sup. Pisa* (3) 22(1968), 571-597; *errata, ibid.* (3). 22: 673, 1968.

[298] C.Sparrow. *The Lorenz equations: bifurcations, chaos, and strange attractors,* volume 41 *of Applied Mathematical Sciences.* Springer-Verlag, New York, 1982.

[299] W.J.Stewart. *Introduction to the numerical solution of Markov chains.* Princeton University Press, Princeton, NJ, 1994.

[300] R.L.Stratonovich. *Topics in the theory of random noise. Vol.* I. Revised English edition. Translated from the Russian by Richard A. Silverman. Gordon and Breach Science Publishers, New York, 1963.

[301] R.L.Stratonovich. *Topics in the theory of random noise. Vol.* II. Revised English edition. Translated from the Russian by Richard A. Silverman. Gordon and Breach Science Publishers, New York, 1967.

[302] D.W.Stroock. *Probability theory, an analytic view.* Cambridge University Press, Cambridge, 1993.

[303] D.W.Stroock. *An introduction to Markov processes* volume 230 of *Graduate Texts in Mathematics.* Springer-Verlag, Berlin, 2005.

[304] D.W.Stroock and S-R.S.Varadhan. *Multidimensional diffusion processes.* Classics in Mathematics. Springer-Verlag, Berlin, 2006.

[305] A.M.Stuart and A.R.Humphries. *Dynamical systems and numerical analysis,* volume 2 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press, Cambridge, 1996.

[306] F.Takens. Motion under the influence of a strong constraining force. In *Global theory of dynamical systems* (*Proc. Internal. Conf., Northwestern Univ., Evanston, Ill,,* 1979). volume 819 *of Lecture Notes in Math.,* pages 425-45. Springer, Berlin, 1980.

[307] L.Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Walt Symposium, Vol.* IV, volume 39 of *Res. Notes in Math.,* pages 136-212. Pitman, Boston, 1979.

[308] L.Tartar. Cours Peccot au College de France. Unpublished, 1989.

[309] L.Tartar. *Nonlocal effects induced by homogenization,* in *PDE and calculus of variations,* F. Culumbini el al., eds. Boston, Birkhauser, 1989.

[310] L.Tartar, H-measures, a new approach for studying homogenization, oscillations and concentration effects in partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A,* 115(3-4): 193-230, 1990.

[311] L.Tartar. An introduction to the homogenization method in optimal design. In *Optimal shape design* (*Troia,* 1998), volume 1740 of *Lecture Notes in Math.,* pages 47-156. Springer, Berlin, 2000.

[312] T.Tassa. Homogenization of two dimensional linear flows with integral invariance. *SIAM J. Appl. Math.*, 57(5): 1390-1405, 1997.

[313] G.I.Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. R. Soc. Lond. A*, 219: 186-203, 1953.

[314] R.Temam. *Infinite-dimensional dynamical systems in mechanics and physics,* volume 68 *of Applied Mathematical Sciences.* Springer-Verlag, New York, 1997.

[315] R.Temam. *Navier-Stokes equations.* AMS Chelsea Publishing, Providence. RI, 2001.

[316] S.Teufel. *Adiabatic perturbation theory in quantum dynamics,* volume 1821 *of Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2003.

[317] A.N.Tikhonov. A.B.Vasil'eva and A.G.Svcshnikov. *Differential equations.* Springer Scries in Soviet Mathematics. Springer-Verlag, Berlin, 1985.

[318] W.Tucker. The Lorenz attractor exists. *C R. Acad. Sci. Paris Ser. I Math.*, 328(12): 1197-1202, 1999.

[319] W.Tucker. A rigorous ODE solver and Smale's 14th problem. *Found. Comput Math.*, 2(1): 53-117, 2002.

[320] N.G.van Kampen. *Stochastic processes in physics and chemistry.* North-Holland Publishing Co., Amsterdam. 1981. Lecture Notes in Mathematics, 888.

[321] N.G.van Kampen. Elimination of fasi variables. *Phys. Rep.*, 124(2): 69-160, 1985.

[322] E.Vanden-Eijnden. Numerical techniques for multi-scale dynamical systems with stochastic effects. *Cammum. Math. Sci.*, 1(2): 385-91, 2003.

[323] M.Vergassola and M.Avellaneda. Scalar transport in compressible flow. *Phys, D,* 106(1-2): 148-66, 1997.

[324] K Verhulst. *Methods and applications of singular perturbations,* volume 50 of *Texts in Applied Mathematics.* Springer, New York, 2005. Boundary layers and multiple timescale dynamics.

[325] D.G.Vlachos. A review of multiscale analysis: Examples from systems biology, materials engineering, and other fluid-surface interacting systems. *Adv. Chem. Eng.*, 30: 1-61, 2005.

[326] J.Walters and Ch.Schuette. Conditional averaging tor diffusive fast-slow systems: a sketch for derivation. In *Analysis, Modeling, and Simulation of Multiscale Problems.* pases 647-82. Springer, 2006.

[327] P.Walters. An *introduction to ergodic theory.* volume 79 *of Graduate Texts in Mathemtlics.* Springer-Verlag, New York, 1982.

[328] W.Wang. D.Cao and J.Duan Effective macroscopic dynamics of stochastic partial differential equations in perforated domains. *SIAM J Math. Anal.*, 38: 1508-1527, 2007.

[329] W.Wang and J Duan. Invariant manifold reduction and bifurcation for stochastic partial differential equations, preprint, 2006.

[330] J-C.Wells. Invariant manifolds on nonlinear operators. *Pacific J. Math.*, 6211 1: 285-93. 1976.

[331] S-Wiggins. *Introduction to applied nonlinear dynamical systems and chaos,* volume 2 of *Texts in Applied Mathematics.* Springer-Verlag, New York. 2003.

[332] E.Wong and M.Zakai. On the convergence of ordinary integrals to stochastic integrals. Ann. *Math. Statist.,* 36: 1560-4, 1965.

[333] J Xin. FroT propagation in heterogeneous media. *SIAM Review,* 42: 161-230, 2000.

[334] D.Xiu and I.G.Kevrekidis. Equation-free, multiscale computation for unsteady random diffusion. *Multiscale Model. Simul.,* 4(3): 915-35 (electronic), 2005.

[335] G.G.Yin and Q.Zhang. *Discrete-time Markov chains,* volume 55 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 2005.

[336] G.G Yin and Q.Zhang. *Continuous-time Markov chains and applications,* volume 37 *of Applications of Mathematics (New York),* Springer-Verlag, New York, 1998.

[337] K.Yosida. *Functional Analysis.* Springer-Verlag, Berlin-Heidelberg, 1995.

[338] L.-S.Young, What are SRB measures and which dynamical systems have them? *J Stat. Phys.,* 108: 733-54, 2002.

[339] E.Zeidler. *Applied Functional Analysis.* volume 108 *of Applied Mathematical Sciences.* Springer-Verlag, New York, 1995.

[340] Ya.B. Zel'dovich-Exact solution of the problem of diffusion in a periodic velocity field, and turbulent diffusion. *Sov. Phys. Dotl.,* 27(10): 797-9, 1982.

[341] V.V.Zhikov. On two-scale convergence. *Tr. Semin. im. I. G. Petrovskogo,* (23): 149-87, 410, 2003.

[342] R.Zwanzig. Diffusion in a rough potential. *Proc. Natl. Acad. Sci. USA,* 85: 2029-30, 1988.

《现代数学译丛》已出版书目

（按出版时间排序）

1 椭圆曲线及其在密码学中的应用——导引 2007. 12 〔德〕Andreas Enge 著 吴 铤 董军武 王明强 译
2 金融数学引论——从风险管理到期权定价 2008. 1 〔美〕Steven Roman 著 邓欣雨 译
3 现代非参数统计 2008. 5 〔美〕Larry Wasserman 著 吴喜之 译
4 最优化问题的扰动分析 2008. 6 〔法〕J. Frédéric Bonnans 〔美〕Alexander Shapiro 著 张立卫 译
5 统计学完全教程 2008. 6 〔美〕Larry Wasserman 著 张 波 等译
6 应用偏微分方程 2008. 7 〔英〕John Ockendon, Sam Howison, Andrew Lacey & Alexander Movchan 著 谭永基 程 晋 蔡志杰 译
7 有向图的理论、算法及其应用 2009. 1 〔丹〕J. 邦詹森 〔英〕G. 古廷 著 姚兵 张忠辅 译
8 微分方程的对称与积分方法 2009.1〔加〕乔治 W. 布卢曼 斯蒂芬 C. 安科 著 闫振亚 译
9 动力系统入门教程及最新发展概述 2009.8 〔美〕Boris Hasselblatt & Anatole Katok 著 朱玉峻 郑宏文 张金莲 阎欣华 译 胡虎翼 校
10 调和分析基础教程 2009.10 〔德〕Anton Deitmar 著 丁勇 译
11 应用分支理论基础 2009. 12 〔俄〕尤里·阿·库兹涅佐夫 著 金成桴 译
12 多尺度计算方法——均匀化及平均化 2010. 6 Grigorios A. Pavliotis, Andrew M. Stuart 著 郑健龙 李友云 钱国平 译

(O−3897.0101)

销售分类建议：高等数学



ISBN 978-7-03-027512-7

9 787030 275127 >

定价：56.00 元